



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12N 15/13, C07K 16/46, 16/00, C12N 5/10, A61K 39/395, G01N 33/577</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 97/38102</b>  <b>(43) International Publication Date:</b> 16 October 1997 (16.10.97)
<b>(21) International Application Number:</b> PCT/EP97/01609  <b>(22) International Filing Date:</b> 26 March 1997 (26.03.97)  <b>(30) Priority Data:</b> 96302412.0 4 April 1996 (04.04.96) EP <b>(34) Countries for which the regional or international application was filed:</b> AT et al.  <b>(71) Applicant (for AU BB CA GB IE IL KE LC LK LS MN MW NZ SD SG SZ TT UG only):</b> UNILEVER PLC [GB/GB]; Unilever House, Blackfriars, London EC4P 4BQ (GB).  <b>(71) Applicant (for all designated States except AU BB CA GB IE IL KE LC LK LS MN MW NZ SD SG SZ TT UG):</b> UNILEVER N.V. [NL/NL]; Weena 455, NL-3013 AL Rotterdam (NL).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> DAVIS, Paul, James [GB/GB]; The Hawthorns, Pavenham Road, Fehmersham, Bedford MK43 7EX (GB). VAN DER LOGT, Cornelis, Paul, Erik [NL/GB]; 1 Bluebell Rise, Peverel Manor Estate, Rushden, Northampton NN10 0TU (GB). VERHOEYEN, Martine, Elisa [BE/GB]; 1 Tintagel Close, Manor Farm Estate, Rushden, Northampton NN10 0QN (GB).	<b>(74) Agent:</b> EVANS, Jacqueline, Gail, Victoria; Unilever plc, Patent Division, Colworth House, Sharnbrook, Bedford MK44 1LQ (GB).  <b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
<b>(54) Title:</b> MULTIVALENT AND MULTISPECIFIC ANTIGEN-BINDING PROTEIN		
<b>(57) Abstract</b>  A multivalent antigen-binding protein comprises a first polypeptide comprising, in series, three or more variable domains of an antibody heavy chain and a second polypeptide comprising, in series, three or more variable domains of an antibody light chain, said first and second polypeptides being linked by association of the respective heavy chain and light chain variable domains, each associated variable domain pair forming an antigen binding site. Methods for their production and uses thereof, in particular for therapeutic and diagnostic applications, are disclosed.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

- 1 -

**MULTIVALENT AND MULTISPECIFIC ANTIGEN-BINDING PROTEIN****FIELD OF THE INVENTION**

5 The present invention relates to multivalent and  
multispecific antigen binding proteins, methods for their  
production and uses thereof. In particular, the  
invention relates to binding proteins comprising  
polypeptides which associate to form multivalent or  
10 multispecific multimers.

**BACKGROUND OF THE INVENTION**

Antibodies are protein molecules having a structure based  
15 on a unit comprising four polypeptides, two identical  
heavy chains and two identical light chains, which are  
covalently linked together by disulphide bonds. Each of  
these chains is folded in discrete domains. The C-  
terminal regions of both heavy and light chains are  
20 conserved in sequence and are called the constant  
regions, comprising one or more so-called C-domains. The  
N-terminal regions of the heavy and light chains, also  
known as V-domains, are variable in sequence and  
determine the specificity of the antibody. The regions  
25 in the variable domains of the light and heavy chains ( $V_L$   
and  $V_H$  respectively) responsible for antigen binding  
activity are known as the hypervariable or  
complementarity determining regions (CDR). Natural  
antibodies have at least two identical antigen-binding  
30 sites defined by the association of the heavy and light  
chain variable regions.

It is known that proteolytic digestion of an antibody can  
lead to the production of antibody fragments. Such  
35 fragments, or portions, of the whole antibody can exhibit  
antigen binding activity. An example of a binding  
fragment is an  $F_{ab}$  fragment which comprises a light chain

- 2 -

associated with the  $V_H$  and  $C_{H1}$  domains of a heavy chain. The bivalent  $F(ab')_2$  fragment comprises two such  $F_{ab}$  fragments connected together via the hinge region, giving two antigen binding sites.  $F_v$  fragments, consisting only of the V-domains of the heavy and light chains associated with each other may also be obtained. These  $F_v$  fragments are monovalent for antigen binding. Smaller fragments such as individual V-domains (domain antibodies or dABs, Ward et al Nature, 341, 544 (1989) and individual CDR's (Williams et al, Proc. Natl. Acad. Sci, USA, 86, 5537 (1989)) have also been shown to retain the binding characteristics of the parent antibody although generally most naturally occurring antibodies need both a  $V_H$  and  $V_L$  to retain full immunoreactivity.

Antibody fragments comprising  $V_H$  and  $V_L$  domains associated together to have antigen binding activity have also been described. The single chain  $F_v$  fragment (scFv) comprises a  $V_H$  domain linked to a  $V_L$  domain by a flexible polypeptide linker such that the domains can associate to form an antigen binding site (see, for example, EP-B-0281604, Bnzon Labs Inc).

Microbial expression systems for producing active antibody fragments are known in the literature. The production of Fab in various hosts such as *E.coli*. (Better et al, Science, 240, 104, (1988)), yeast (Horwitz et al, Proc. Natl. Acad. Sci, USA, 85, 8678 (1988)) and the filamentous fungus *Trichoderma reesei* (Nyyssönen et al, Bio/Technology, 11, 591 (1993)) have previously been described, for example. It is also known that plants can be used as hosts for the production of SCFv fragments (Owen et al, Bio/Technology, 10, 790 (1992)) as well as whole antibodies.

An advantage of using antibody fragments rather than whole antibodies in diagnosis and therapy lies in their

- 3 -

smaller size. They are likely to be less immunogenic than whole antibodies and more able to penetrate tissue. A disadvantage associated with the use of fragments such as the  $F_{ab}$ ,  $F_v$ , and  $S_cF_v$  antibody fragments described above, however is that they have only one binding site for antigen binding as compared to the two or more sites contained in the whole antibody, preventing polyvalent binding to the antigen and hence leading to reduced avidity.

In an attempt to overcome this problem, attention has been directed to providing multivalent antigen binding proteins, that is binding proteins having more than one antigen binding site. In addition, there has been interest in producing antigen-binding proteins having multiple specificities capable of binding to different antigenic determinants and containing antigen binding domains derived from different sources. Antigen-binding proteins having distinct binding specificities may be useful, for example, in targeting effector cells to target cells by virtue of the specific binding of the different binding domains. By way of illustration, a bispecific antigen binding protein having specificity for both tumour cells and cytotoxic drugs may be used to target specifically cytotoxic drug to tumour cell in an efficient manner. By avoiding the need for chemical modification, adverse immune responses may be avoided.

Hitherto, the potential application of multivalent and multispecific antigen binding proteins have been hindered by the difficulties in generating and purifying such molecules.

Recombinant antigen-binding proteins having two binding sites may be prepared by methods such as chemical cross-linking of cysteine residues, either through cysteine residues introduced at the C-terminus of the  $V_H$  of an  $F_v$

- 4 -

(Cumber et al, J.Immunol., 149, 120 (1992)), through the hinge cysteine residues in F<sub>ab</sub> to generate (Fab<sup>1</sup>), (Carter et al, Bio/Tech., 10, 163 (1992)) or at the C-terminus of the V<sub>L</sub> of an scFv (Pack and Plückthun, Biochemistry, 31, 1579 (1992)). Alternatively, the production of bivalent and bispecific antibody fragments based on the inclusion of F<sub>ab</sub> fragments of C-terminal peptide sequences which promote dimerisation has been described. (Kostelny et al, J.Immunol., 148, 1547).

Bivalent or bispecific antibody fragments comprising a binding complex containing two polypeptide chains, one comprising two heavy chain variable domains (V<sub>H</sub>) in series and the other comprising two light chain variable domains (V<sub>L</sub>) in series are described in our pending European Patent Application No. 95307332.7.

Multivalent and/or multispecific antibody fragments are described in WO 94/09131 (Scotgen Limited). Specific binding proteins having two binding regions, contained at least in part on first and second polypeptide chains which chains additionally incorporate associating domains capable of binding to each other causing the polypeptide chains to combine are disclosed therein. It is disclosed that the first and second binding regions preferably are antibody antigen-binding domains, for example comprising V<sub>H</sub> and V<sub>L</sub> regions contained in a Fab fragment or in a single-chain Fv fragment, or may be derived from just one of the V<sub>H</sub> or V<sub>L</sub> regions of an antibody. The associating domains may suitably be derived from an antibody and may be inter alia antibody V<sub>H</sub> and V<sub>L</sub> regions. It is further disclosed that using a V<sub>H</sub>/V<sub>L</sub> domain combination to achieve association leads to the creation of a supplementary Fv domain such that the antibody produced may be trivalent. Schematic representations of the arrangements suggested in WO 94/09131 to produce trivalent fragments are shown in Figure 1A.

- 5 -

WO 93/11161 (Bnzon Inc) describes multivalent antigen-binding proteins comprising two or more single-chain protein molecules, each single chain molecule comprising first and second polypeptides each comprising the binding portion of the variable region of an antibody heavy or light chain with the polypeptides being linked together via a peptide linker. Hypothetical trimers and tetramers are discussed, comprising three or four single-chain antigen binding proteins as appropriate. Schematic representations of the trivalent arrangements suggested are shown in Figure 1B.

WO 91/19739 (Celltech Limited) discloses multivalent antigen binding proteins comprising an Fv fragment bound to at least one further Fv fragment by a connecting structure which links the Fv fragments together but which maintains them spaced apart such that they can bind to adjacent antigenic determinants. Conveniently the connecting structure consists of a spacing polypeptide and a linkage unit such as a cross-linking maleimide linker or a molecule which allows for non-covalent binding. Particularly preferred connecting structures which are disclosed are based on antibody joining and hinge region sequences.

#### SUMMARY OF THE INVENTION

According to the present invention there is provided a multivalent antigen binding protein comprising:

- a first polypeptide comprising in series, three or more variable domains of an antibody heavy chain; and
- a second polypeptide comprising, in series, three or more variable domains of an antibody light chain,

- 6 -

said first and second polypeptides being linked by association of the respective heavy chain and light chain variable domains, each associated variable domain pair forming an antigen binding site.

5

As used herein, the term multivalent means more than one antigen binding site.

10

Preferably the first polypeptide comprises three variable domains of an antibody heavy chain and the second polypeptide comprises three variable domains of an antibody light chain, providing a trivalent protein.

15

It will be appreciated that the polypeptides may comprise heavy or light chains, variable domains, as appropriate, or functional equivalents thereof.

20

The respective heavy or light chain variable domains may suitably be linked without any intervening linker. According to a preferred embodiment, however, the variable domains contained in the individual polypeptides are linked by peptide linkers. Preferably the peptide linker is flexible, allowing the variable domains to flex in relation to each other such that they can bind to multiple antigenic determinants simultaneously. It will be appreciated that the binding of the linker to the individual heavy or light chain variable domains will be such that it does not affect the binding capacity of the binding site formed by the associated variable domain pair. Conveniently the peptide linker comprises from 16 to 19 amino acid residues. A preferred, peptide linker for heavy chain domains is (Gly,Ser),AlaGlySerAla and for the light chain domains is (Gly,Ser),Val.

30

35

It will be appreciated that if two or more of the associated variable domain pairs ( $V_H/V_L$  pairs) have the same antigen specificity, for example if they are derived



- 7 -

from the same parent antibody or fragment thereof or from different antibodies which bind the same epitope, then a binding protein which binds more than one molecule of the same type will be produced.

5

According to one embodiment, where the binding protein according to the invention comprises three antigen binding sites which are able to bind different epitopes from each other, a trivalent trispecific protein is produced.

10

In another embodiment, where the binding protein according to the invention comprises three associated variable domain pair binding sites, two of which sites bind the same epitopes, a trivalent, bispecific protein is provided. Where all three binding sites have the same antigen specificity, a trivalent, monospecific binding protein is provided.

15

The invention also provides nucleotide sequences coding for the polypeptides of the multivalent antigen binding protein according to the invention and cloning and expression vectors containing such nucleotide sequences.

20

The invention further provides host cells transformed with vectors containing such nucleotide sequences and methods of producing such polypeptides by expression of the nucleotide sequences in such hosts.

25

The invention further provides a process for preparing a multivalent antigen binding protein as set forth above comprising:

30

(i) transforming one or more hosts by incorporating genes encoding said first and second polypeptides;

35

(ii) expressing said genes in said host or hosts;

- 8 -

- (iii) allowing said first and second polypeptides to combine to form the antigen binding protein.

Suitably the host or hosts may be selected from  
5 prokaryotic bacteria, such as Gram-negative bacteria, for example *E.Coli*, and Gram-positive bacteria, for example *B. subtilis* or lactic acid bacteria, lower eukaryotes such as yeasts, for example belonging to the genera *Saccharomyces Kluyveromyces* or *Trichoderma*, moulds such  
10 as those belonging to the genera *Aspergillus* and *Neurospora* and higher eukaroytes, such as plants, for example tobacco, and animal cells, examples of which are myeloma cells and CHO, COS cells and insect cells. A particularly preferred host for use in connection with  
15 the present invention is COS (monkey kidney) cells.

Techniques for synthesising genes, incorporating them into hosts and expressing genes in hosts are well known in the art and the skilled person would readily be able  
20 to put the invention into effect using common general knowledge. Proteins according to the invention may be recovered and purified using conventional techniques such as affinity chromatography, ion exchange chromatography or gel filtration chromatography.

25 The activity of the multivalent binding proteins according to the invention may conveniently be measured by standard techniques known in the art such as enzyme-linked immunosorbant assay (ELISA), radioimmune assay  
30 (RIA) or by using biosensors.

The multivalent antigen binding proteins of the present invention may suitably be used in diagnostics or therapy for example in targeting a tumour cell with natural  
35 killer cells and cytotoxic agent. Other uses for which the multivalent binding proteins according to the invention are useful include those uses for which

- 9 -

antibodies or fragments thereof are commonly used, including for immunoassays and in purification.

According to a particular preferred embodiment, multi-enzyme complexes may be assembled, at a target, for example a cell surface. As an illustration, multivalent binding proteins according to the invention may be used to target cell killing enzymes such as an oxidase (for example glucose oxidase) and peroxidase (for example horseradish peroxidase) to a target species which is an antigenic component of dental plaque, such as *S. sanguis* or *S. mutans*. Complexes comprising enzyme, coenzyme and target antigen may also conveniently be assembled.

Accordingly, the invention also provides compositions comprising the multivalent antigen binding proteins according to the invention, conveniently in combination with a cosmetically or pharmaceutically acceptable carrier, diluent or excipient. Methods of treatment using the multivalent antigen binding proteins according to the invention are also provided.

For use in diagnosis or therapy, the multivalent antigen binding proteins according to the invention may conveniently be attached to an appropriate diagnostically or therapeutically effective agent or carrier by methods conventional in the art.

An advantage of using multivalents antigen binding proteins according to the invention over multivalent binding proteins prepared by existing techniques known in the art is that the "self-assembling" association of the respective heavy and light chain variable domains to form the multivalent binding sites avoids the need for chemical coupling steps or the introduction of linking residues to stabilise the multivalent constructs, thereby minimising the risk of eliciting an immune response to such molecules when the resulting multivalent binding

- 10 -

proteins are used in therapy.

5 A particular advantage of molecules according to the present invention is that they may conveniently be purified straight from the supernatant using conventional purification techniques. As they are self-assembling, there is no need to purify individual subunits prior to coupling as in existing techniques.

10 The present invention may be more fully understood with reference to the following description, when read together with the accompanying drawings.

#### BRIEF DESCRIPTION OF THE DRAWINGS

15

Figures 1A and 1B show schematic representations of published arrangements of heavy and light chain V-domain gene fragments that have been suggested to produce trispecific or trivalent antibody fragments:

20

- |       |  |             |
|-------|--|-------------|
| A)    | scFv1-VLa + scFv2-VHa (2 chains)         | WO 94/09131 |
| B)    | Fab1-VLa + Fab2-VHa (4 chains)           | WO 94/09131 |
| C)    | scFv1-VLa-CLa + scFv1-VHa-CHa (2 chains) | WO 94/09131 |
| D)    | Fab1-VLa-CLa + Fab2-VHa-CHa (4 chains)   | WO 94/09131 |
| 25 E) | scFv1 + scFv2 + scFv3 (3 chains)         | WO 93/11161 |
| F)    | VH1-VL2 + VH2-VL3 + VH3-VH1 (3 chains)   | WO 93/11161 |

30 Figure 2A/B shows the nucleotide sequence of the EcoRI-HindIII insert of pGOSA.E2t containing DNA encoding pelB leader-VH4715-linker-VL3418 and DNA encoding pelB leader-VL3418-linker-VH4715-hydrophil2 tag (SEQ ID No. 1).

#### Figure 3

35 A) shows the nucleotide sequence of the HindIII-EcoRI insert of plasmid scFv.Lys with DNA encoding pelB leader-VHLys-linker-VLLys (SEQ ID No. 2).

- 11 -

B) shows the nucleotide sequence of the HindIII-EcoRI insert of plasmid scFv.4715.2t with DNA encoding pelB leader-VH4715.2t (SEQ ID No. 3).

5     **Figure 4** shows the nucleotide sequence of the genomic leader sequence of the anti-NP antibody (Jones et al, Nature, 321, 522). Exon sequences are indicated with shaded boxes. NcoI and PstI restriction sites are in bold and underlined (SEQ ID No. 4).

10

**Figure 5** gives a schematic representation of the eukaryotic expression vector pSV.51.

15     **Figure 6** gives an overview of the pUC19 double head (A) and triple head (B) constructs. The position of the oligonucleotides and the restriction sites used for assembling double and triplehead pUC constructs are indicated.

20

**Figure 7**

A) shows the origin of the VH-C-linker and VL-C-linker fragments.

25

B) gives a schematic representation of the construction of the pUC.19-triple-head vectors.

**Figure 8**

30

A) gives a schematic representation of the construction of the Euka.VH and Euka.VL vectors.

B) gives a schematic representation of the construction of the pSV.VH expression vectors.

35

C) gives a schematic representation of the construction of the pSV.VL expression vectors.

- 12 -

Figure 9 shows the expression of the trispecific Golysan proteins on an SDS-PAGE gel containing total COS culture supernatant. Crude supernatants of COS cells transfected with pSV expression vectors were separated on SDS-PAGE gels. The proteins were transferred onto a nitrocellulose membrane and the VH3 and VL3-2t were detected using anti-VH and anti-hydrophil 2 tag specific monoclonal antibodies respectively. (A=anti-Hydro-II, B=anti-Hydro-II + anti-VH) Samples: M) Low Molecular Weight Markers, 1) pSV.K + pSV.V, 2) pSV.K + pSV.W, 3) pSV.M + pSV.V, 4) pSV.M + pSV.W.

Figure 10 shows the results of three ELISA's. Lysozyme, Glucose oxidase and S.sanguis binding activity was determined in crude COS supernatants by measuring 1) Lysozyme-Glucose oxidase (=LYSOX), 2) Glucose oxidase-S.sanguis (=GOSA) and 3) Lysozyme-S.sanguis (=LYSAN) bispecific binding activities.

Figure 11 shows the results of three ELISA's. Lysozyme, Glucose oxidase and S.sanguis binding activity of purified Golysan.A (A) and Golysan.B (B) was determined by measuring 1) Lysozyme-Glucose oxidase (=LYSOX), 2) Glucose oxidase-S.sanguis (=GOSA) and 3) Lysozyme-S.sanguis (=LYSAN) bispecific binding activities.

Figure 12 shows the nucleotide sequence of the EcoRI-HindIII insert of pUR.4124 containing DNA (see SEQ ID NO: 23) encoding V<sub>L</sub>Lys-Linker-V<sub>H</sub>Lys.

Figure 13 shows the nucleotide sequence of the HindIII-EcoRI insert of plasmid Fv.3418 (see SEQ ID NO: 24) containing DNA encoding pelB leader-V<sub>H</sub>3418 and pelB leader-V<sub>L</sub>3418.

Figure 14 shows the nucleotide sequence of the HindIII-EcoRI insert of plasmid Fv.4715-myc (see SEQ ID NO: 25)

- 13 -

containing DNA encoding pelB leader-V<sub>H</sub>4715 and pelB leader-V<sub>L</sub>4715-Myc tag.

5      **Figure 15** shows the nucleotide sequence of the *Hind*III-*Eco*RI insert of scFv.4715-myc containing DNA (see SEQ ID NO: 26) encoding pelB leader-V<sub>H</sub>4715-Linker-V<sub>L</sub>4715-Myc tag.

10      **Figure 16a/b** shows the nucleotide sequence of the *Hind*III-*Eco*RI insert of pGOSA.E (see SEQ ID NO: 27) containing DNA encoding pelB leader-V<sub>H</sub>4715-Linker-V<sub>L</sub>3418 and pelB leader-V<sub>L</sub>3418-Linker-V<sub>H</sub>4715.

15      **Figure 16c** gives an overview of the oligonucleotides and their positions in pGOSA.E that can be used to replace V-domain gene fragments.

**Figure 17** shows the construction of plasmid pGOSA.A.

20      **Figure 18** shows the construction of plasmid pGOSA.B.

**Figure 19** shows the construction of plasmid pGOSA.C.

**Figure 20** shows the construction of plasmid pGOSA.D.

25      **Figure 21** shows the construction of plasmid pGOSA.E.

**Figure 22** shows the source of fragment PCR.I *Bst*EII/*Sac*I.

30      **Figure 23** shows the source of fragment PCR.IV *Xho*I/*Eco*RI.

**Figure 24** shows the source of fragment PCR.V *Sal*I/*Eco*RI.

**Figure 25** shows the source of fragment PCR.III *Nhe*I/*Sac*I.

35      **Figure 26** shows the source of fragment PCR.II *Sfi*I/*Eco*RI.

**Table 1** shows the nucleotide sequence of all

- 14 -

oligonucleotides used in the construction of the described double and triple head constructs.

5 **Table 2** lists all pSV expression constructs described in this specification.

The following examples are provided by way of illustration only:

10 **EXAMPLES**

**General Experimental**

**Strains, Plasmids and Media**

15 All cloning steps were performed in E.Coli JM109 or E.Coli XL-1 Blue. Cultures were grown in 2xTY/Amp/Glucose medium (16g tryptone, 10g yeast extract, 5g NaCl per liter H<sub>2</sub>O supplemented with 2% glucose and 100µg/ml ampicillin). Transformations were plated out on  
20 SOBAG plates (20g tryptone, 5g yeast extract, 15g agar, 0.5g NaCl per liter H<sub>2</sub>O plus 10mM MgCl<sub>2</sub>, 2% glucose, 100µg/ml ampicillin). The bicistronic E.coli vectors used are derivatives of pUC19. The COS expression vector pSV.51 (LMBP strain nr 1829) was obtained from the LMBP  
25 Culture collection (Laboratory of Molecular Biology University Gent). COS-1 cells (ECACC No: 88031701; African green monkey kidney cells) were obtained from the European Collection of Animal Cell Cultures (ECACC). All tissue culture reagents were from Gibco BRL (Life  
30 Technologies, Paisley, UK)

**DNA Manipulations**

**Oligonucleotides and PCR**

35 The oligonucleotide primers used in the PCR reactions were synthesized on an Applied Biosystems 381A DNA Synthesiser by the phosphoramidite method. The primary structures of the oligonucleotide primers used in the



- 15 -

construction of the trispecific pSV constructs (Table 2) are shown in Table 1. Reaction mixture used for amplification of DNA fragments were 10mM Tris-HCl, pH8.3, 2.5mM MgCl<sub>2</sub>, 50mM KCl, 0.01% gelatin (w/v), 0.1% Triton X-100, 400mM of each dNTP, 5.0 units of Vent DNA polymerase (New England Biolabs), 100ng of template DNA, and 500ng of each primer (for 100μl reactions). Reaction conditions were: 94°C for 4 minutes, followed by 33 cycles of each 1 minute at 94°C, 1 minute at 55°C, and 1 minute 72°C.

**Plasmid DNA\Vector\Insert preparation and ligation\transformation.**

Plasmid DNA was prepared using the 'Qiagen P-100 and P-500 Midi/Maxi-DNA Preparation' system. Vectors and inserts were prepared by digestion of 10μg (for vector preparation) or 20μg (for insert preparation) with the specified restriction endonucleases under appropriate conditions (buffers and temperatures as specified by suppliers). Klenow fill-in reactions and dephosphorylation with Calf Intestine Phosphorylase were performed according to the manufacturers instructions. Vector DNA's and inserts were separated through agarose gel electrophoresis and purified with DEAE-membranes NA45 (Schleicher & Schnell) as described by Maniatis et al. (Molecular cloning: a Laboratory manual, Cold Spring Harbour, N.Y. (1982)) Ligations were performed in 20μl volumes containing 30mM Tris-HCl pH7.8, 10mM MgCl<sub>2</sub>, 10mM DTT, 1mM ATP, 300-400ng vector DNA, 100-200ng insert DNA and 1 Weiss unit T<sub>4</sub> DNA ligase. After ligation for 2-4 h at room temperature, CaCl<sub>2</sub> competent E. coli JM109 or XL-1 Blue (Maniatis et al) were transformed using 7.5μl ligation reaction. The transformation mixtures were plated onto SOBAG plates and grown overnight at 37°C. Correct clones were identified by restriction analysis and verified by automated dideoxy sequencing (Applied Biosystems).

- 16 -

**Restriction digestion of PCR products**

Following amplification each reaction was checked for the presence of a band of the appropriate size by agarose gel electrophoresis. One or two 100 $\mu$ l PCR reaction mixtures of each of the PCR reactions, together containing approximately 2-4 $\mu$ g DNA product were subjected to phenol-chloroform extraction, chloroform extraction and ethanol precipitation. The DNA pellets were washed twice with 70% ethanol and allowed to dry. Next, the PCR products were digested overnight (18 h) in 200 $\mu$ L 1xBuffer with excess of the appropriate restriction enzyme.

**Transformation of COS Cells**

Cos-1 cells were maintained in DMEM culture medium with glutamine (2mM), Penicillin (100U/mL), streptomycin (100 $\mu$ g/mL) containing 10% F.C.S. For transient transfection assays 1-3 x10<sup>5</sup> COS-1 cells were seeded in 3 cm-diameter tissue culture dishes (2mL). The cells were incubated at 37°C in a CO<sub>2</sub> incubator until cells were 50-80% confluent (overnight). For each transfection the following mixes were prepared: A) 1 $\mu$ g of each of the specified DNA's in 100 $\mu$ L Opti-MEM-I Reduced Serum Medium, B) 1 $\mu$ L LipofectAmine in 100 $\mu$ L Opti-MEM-I Reduced Serum Medium. Mixes A and B were combined (gently). After allowing the DNA-liposome complexes to form for 30-45 minutes at room temperature, 0.8mL Opti-MEM-I Reduced Serum Medium was added to each lipid DNA complex containing tube. The COS-1 cells were washed once with 2mL of Opti-MEM-I Reduced Serum Medium and overlaid with the diluted complex solution. The COS-1 cells were incubated for 5 hr at 37°C. Following incubation, 2mL growth medium was added. 20 hours following transfection the medium was replaced with 2mL fresh growth medium containing 0.1mM Na-butyrate. After 48 hours incubation at 37°C the supernatant was harvested and assayed for the presence of antibody fragments.

- 17 -

**ELISA****A) GOSA: Glucose Oxidase and *S.sanguis* binding activity**

96 well ELISA plates (Greiner HC plates) were activated  
overnight at 37°C with 200µl/well of a 1/10 dilution of  
an overnight culture of *Streptococcus sanguis* cells in  
0.05M sodium carbonate buffer pH9.5 was used to sensitise  
each well. Following one wash with PBST, the antigen  
sensitised plates were pre-blocked for 1 hour at 37°C  
with 200µl/well blocking buffer (1% BSA, 0.15% Tween in  
PBS). 50µl COS culture supernatants (neat or diluted  
with PBS) plus 50µl blocking buffer containing glucose  
oxidase (50µg/ml) was added to the *Streptococcus Sanguis*  
sensitised plate and incubated for 2 hours at 37°C.  
Following 4 washes with PBS-T, bound glucose oxidase was  
detected by adding 100µl substrate to each well (70mM Na-  
citrate, 320mM Na-phosphate, 27mg/ml glucose, 0.5µg/ml  
HRP, 100µg/ml TMB). The colour reaction was stopped  
after 1 hour by the addition of 35µl 2M HCl and the A450  
was measured.

**B) LYSOX: Lysozyme and Glucose Oxidase binding activity**

96 well ELISA plates (Greiner HC plates) were activated  
overnight at 37°C with lysozyme (50µg/mL in 0.05M sodium  
carbonate buffer pH9.5; 200µl/well). Following one wash  
with PBST, the antigen sensitised plates were pre-blocked  
for 1 hour at 37°C with 200µl/well blocking buffer (1%  
BSA, 0.15% Tween in PBS). 50µl COS culture supernatants  
(neat or diluted with PBS) plus 50µl blocking buffer  
containing glucose oxidase (50µg/ml) was added to the  
*Streptococcus Sanguis* sensitised plate and incubated for  
2 hours at 37°C. Following 4 washes with PBS-T, bound  
glucose oxidase was detected by adding 100µl substrate to  
each well (70mM Na-citrate, 320mM Na-phosphate, 27mg/ml  
glucose, 0.5µg/ml HRP, 100µg/ml TMB). The colour  
reaction was stopped after 1 hour by the addition of 35µl  
2M HCl and the A450 was measured.

- 18 -

**C) LYSAN: *S.sanguis* and Lysozyme binding activity**

96 well ELISA plates (Greiner HC plates) were activated overnight at 37°C with 200µl/well of a 1/10 dilution of an overnight culture of *Streptococcus sanguis* cells in 0.05M sodium carbonate buffer pH9.5 was used to sensitise each well. Following one wash with PBST, the antigen sensitised plates were pre-blocked for 1 hour at 37°C with 200µl/well blocking buffer (1% BSA, 0.15% Tween in PBS). 50µl COS culture supernatants (neat or diluted with PBS) plus 50µl blocking buffer was added to the *Streptococcus Sanguis* sensitised plate and incubated for 2 hours at 37°C. Following 4 washes with PBS-T, 50µL blocking buffer containing Alkaline-Phosphatase conjugated Lysozyme (100µ/mL). Unbound Lysozyme was removed by 4 washes with PBS-T. Bound Lysozyme was detected by adding 100µL substrate solution to each well (1mg/ml pNPP in 1M diethanolamine, 1mM MgCl<sub>2</sub>). After 1 hour the A405 was measured.

**EXAMPLE 1: Construction of the pSV.Golysan expression vectors**

The construction of the pSV COS expression vectors consisted of three stages:

1A): Assembly of 2 heavy chain variable domains and 2 light chain variable domains in a pUC based *E.Coli* expression vector thus constructing the VH<sub>A</sub>-VH<sub>B</sub> and VL<sub>A</sub>-VL<sub>B</sub> modules respectively.

1B): Assembly of 3 heavy chain variable domains and 3 light chain variable domains in a pUC based *E.Coli* expression vector thus constructing the VH<sub>A</sub>-VH<sub>B</sub>-VH<sub>C</sub> and VL<sub>A</sub>-VL<sub>B</sub>-VL<sub>C</sub> modules respectively.

2) Linking the VH<sub>A</sub>-VH<sub>B</sub>, VH<sub>A</sub>-VH<sub>B</sub>-VH<sub>C</sub> and VL<sub>A</sub>-VL<sub>B</sub>, VL<sub>A</sub>-VL<sub>B</sub>-VL<sub>C</sub> to the genomic anti-NP leader sequence in the

- 19 -

intermediate "EUKA" vectors to ensure efficient secretion by COS cells.

- 5        3)    Inserting the leader-VH<sub>A</sub>-VH<sub>B</sub>, leader-VH<sub>A</sub>-VH<sub>B</sub>-VH<sub>C</sub> and  
         leader-VL<sub>A</sub>-VL<sub>B</sub>, leader-VL<sub>A</sub>-VL<sub>B</sub>-VL<sub>C</sub> as XbaI/XbaI  
         fragments downstream of the SV40 promoter in the COS  
         expression vector pSV.51.

***ad.1) E.coli expression vectors.***

- 10       The E.coli expression vectors are derivatives of pUC.19  
         containing a HindIII-EcoRI fragment that in the case of  
         the scFv.lys-myc contains a pelB signal sequence fused to  
         the 5' end of the heavy chain V-domain that is directly  
15       linked to the corresponding light chain V-domain of the  
         antibody through a connecting sequence that codes for a  
         flexible peptide (Gly,Ser), thus generating a single-chain  
         molecule. In the 'double head' expression vector both  
         the heavy chain and the light chain V-domains of the  
         antibody are preceded by a ribosome binding site and a  
20       pelB signal sequence in an artificial dicistronic operon  
         under the control of a single inducible promoter.  
         Expression of these constructs is driven by the inducible  
         lacZ promoter. The nucleotide sequence of the HindIII-  
         EcoRI inserts of the scFv.lys-myc, scFv.4715.2t and  
25       pGOSA.E2t constructs used for the generation of the  
         trispecific antibody fragments are listed in Figures 3  
         and 2 respectively.

***ad.1A) Assembly of bi-specific fragments or double heads.***

- 30       The construct pGOSA.E2t (Figures 2 and 6A) is derived  
         from the E.coli expression construct pGOSA.E. The  
         construction of pGOSA.E has been described in detail in  
         preparation 1 below.
- 35       In contrast with pGOSA.E, pGOSA.E2t contains a peptide  
         tag at the C-terminus of the Variable light chain. Using  
         oligonucleotides DBL3 and DBL.4 the VL4715 gene fragment

- 20 -

was amplified using scFv.4715.2t as a template. The Sall/BamHI VH4715.2t PCR fragment and the Hydrophil-2 tag containing BamHI/EcoRI fragment from scFv.4715.2t (Figure 3B) were used to replace the Sall/EcoRI VH4715 fragment in pGOSA.E thus producing pGOSA.E2t.

The vector pGOSA.E2t and the oligonucleotides in Table 1 have been designed to enable most specificities to be cloned into the pGOSA.E2t construct (Figure 6A). The upstream V<sub>H</sub> domain can be replaced by any PstI-BstEII V<sub>H</sub> gene fragment obtained with oligonucleotides PCR.51 and PCR.89. The oligonucleotides DBL.1 and DBL.2 were designed to introduce SfiI and NheI restriction sites in the V<sub>H</sub> gene fragments thus allowing cloning of those V<sub>H</sub> gene fragments into the SfiI-NheI sites as the downstream V<sub>H</sub> domain. Using this approach the following V<sub>H</sub><sub>A</sub>-V<sub>H</sub><sub>B</sub> combinations were constructed: VH4715-VH3418, VH4715-VHlys, VH3418-VHlys, VHlys-VH3418.

All V<sub>L</sub> gene fragments obtained with oligonucleotides PCR.116 and PCR.90 can be cloned into the position of the 3418 V<sub>L</sub> gene fragment as a SacI-XhoI fragment. A complication here however is the presence of an internal SacI site in the 3418 V<sub>H</sub> gene fragment. Oligonucleotides DBL.3 and DBL.4 are designed to allow cloning of V<sub>L</sub> gene fragments into the position of the 4715 V<sub>L</sub> gene fragment as a Sall-BamHI fragment. A complication here however is the presence of an internal BamHI site in the hydrophil-2-tag gene fragment (2t). Using this approach the following V<sub>L</sub><sub>A</sub>-V<sub>L</sub><sub>B</sub> combinations were constructed: VL3418-VL4715.2t, VLlys-VL4715.2t and VLlys-VL3418.2t.

***ad.1B) Assembly of tri-specific fragments or triple heads.***

Amplification of the VH-linker fragments using either scFv (VH-linker-VL) or bi-specific constructs (VH-linker-VH) as template with the primer combination DBL.1/DBL.5

- 21 -

(Figure 7A) yields one of the building blocks for the construction of the  $VH_A$ - $VH_B$ - $VH_C$  modules. The VH-linker DBL.1/DBL.5 PCR fragment is digested with SfiI and inserted into the SfiI site that is present between the linker sequence and the downstream VH domain in all bi-specific constructs (Figure 7B) thus producing a  $VH_A$ - $VH_B$ - $VH_C$  module. Using this approach the following  $VH_A$ - $VH_B$ - $VH_C$  combinations were constructed for this filing: VH4715-VHlys-VH3418 and VHlys-VH4715-VH3418.

10

~~Using a bi-specific construct (VL-linker-VL) as the template in an amplification reaction with the primer combination DBL.3/DBL.6~~ (Figure 7A) yields the VL-linker building block for the construction of the  $VL_A$ - $VL_B$ - $VL_C$  modules. The VL-linker DBL.3/DBL.6 PCR fragment is digested with SalI and inserted into the SalI site that is present between the linker sequence and the downstream VL domain in all bi-specific constructs (Figure 7B) thus producing a  $VL_A$ - $VL_B$ - $VL_C$  module. Using this approach the following  $VL_A$ - $VL_B$ - $VL_C$  combinations were constructed: VLlys-VL4715-VL3418.2t and VL3418-VLlys-VL4715.2t.

15

20

25

A schematic representation of the final tri-specific constructs is shown in Figure 6B.

**ad.2) Linking the variable region domains to the leader sequence.**

30

35

The HindIII/EcoRI polylinker of pUC19 was replaced with a synthetic EcoRI/HindIII 'Euka' polylinker. This was achieved by annealing and inserting the synthetic oligonucleotides Euka.1 and Euka.2 (Table 1) into EcoRI/HindIII digested pUC19 vector. The resulting Euka.pUC vector contains all restriction sites needed for the subcloning of the leader sequence and the VH and VL domains. The NcoI/PstI genomic anti-NP leader sequence fragment was cloned into the NcoI/PstI digested Euka.pUC vector yielding the Euka.VH construct (Figure 8A).

- 22 -

Oligonucleotides ML.1 and ML.2 (Table 1) were used in an amplification reaction to introduce a SacI site at the 3' end of the leader sequence that allows the construction of leader-VL fusions. The NcoI/SacI leader sequence PCR fragment was inserted into NcoI/SacI digested Euka.pUC vector yielding the Euka.VL construct (Figure 8A).

The  $VH_A$ - $VH_B$  and  $VH_A$ - $VH_B$ - $VH_C$  modules were excised from the pUC expression vectors as PstI/NheI fragments and inserted into PstI/NheI digested Euka.VH vector (Figure 8B). Using this approach the following leader- $VH_A$ - $VH_B$  and leader- $VH_A$ - $VH_B$ - $VH_C$  combinations were constructed for this filing: Euka.B: leader-VH4715-VH3418, Euka.D: leader-VH4715-VHlys, Euka.G: leader-VH3418-VHlys, Euka.K: leader-VH4715-VHlys-VH3418 and Euka.M: leader-VHlys-VH4715-VH3418.

The  $VL_A$ - $VL_B$  and  $VL_A$ - $VL_B$ - $VL_C$  modules were excised from the pUC expression vectors as EcoRI-Klenow/SacI fragments and inserted into NotI-Klenow/SacI treated Euka.VL vector (Figure 8C). Using this approach the following leader- $VL_A$ - $VL_B$  and leader- $VL_A$ - $VL_B$ - $VL_C$  combinations were constructed: Euka.N: leader-VL3418-VL4715.2t, Euka.P: leader-VLlys-VL4715.2t Euka.S: leader-VLlys-VL3418.2t, Euka.V: leader-VLlys-VL4715-VL3418.2t and Euka.W: leader-VL3418-VLlys-VL4715.2t.

**ad.3) Subcloning of leader-variable domain fusions into the pSV.51 expression vector**

All leader- $VH_A$ - $VH_B$ , leader- $VH_A$ - $VH_B$ - $VH_C$ , leader- $VL_A$ - $VL_B$  and leader- $VL_A$ - $VL_B$ - $VL_C$  combinations were excised from the 'Euka' vectors as XbaI/XbaI fragments and subcloned downstream of the SV40 promoter in pSV.51 (Figure 5) by insertion into the XbaI site (Figure 8B and 8C). After confirmation of the correct orientation of the inserts the pSV expression vectors were used to transfect COS-1 cells (see Example 2). The pSV expression vectors used



- 23 -

are listed in Table 2.

**Example 2: Bifunctional binding activity of Golysan triple heads**

5

This example describes the production of three types of bispecific binding activity by COS-1 cells transfected with expression plasmids encoding the corresponding  $VH_A$ - $VH_B$ - $VH_C$  and  $VL_A$ - $VL_B$ - $VL_C$  genes fragments.

10

**1. Production of antibody fragments by COS-1 cells**

Supernatants of COS-1 cells transfected with combinations of pSV- $VH_A$ - $VH_B$ - $VH_C$  and pSV- $VL_A$ - $VL_B$ - $VL_C$  expression plasmids were separated on 10% SDS-PAGE and transferred onto nitrocellulose. The resulting Western blots were screened with a monoclonal antibody recognising a peptide sequence in framework 4 of the VH domains (region encoded by PCR.89: conserved in all used VH domains, {in-house reagent}) and/or a monoclonal specific for the hydrophilic 2 tag. As shown in Figure 9 all supernatants contained products with the expected molecular weight of the  $VH_A$ - $VH_B$ - $VH_C$  and  $VL_A$ - $VL_B$ - $VL_C$  fragments, indicating that the COS cells were successfully transfected and were secreting the produced antibody fragments into the culture medium at detectable levels.

25

**2. Bifunctional binding activity**

Supernatants of COS-1 cells transfected with single pSV expression plasmids and combinations of pSV expression plasmids were tested for the production of bifunctional binding activity using ELISA format:

30

\* Supernatants of COS-1 cells transfected with the bispecific positive controls 'LYSAN' (pSV.D + pSV.P), 'LYSOX' (pSV.G + pSV.S) and 'GOSA' (pSV.B + pSV.N) only produced LYSAN, LYSOX and GOSA bispecific activity respectively (Figure 10). No significant cross

35

- 24 -

reactivity was detected.

5 \* Supernatants of COS-1 cells transfected with only one expression vector encoding either one of the  $VH_A$ - $VH_B$ - $VH_C$  fragments (pSV.K and pSV.M) or one of the  $VL_A$ - $VL_B$ - $VL_C$  fragments (pSV.V and pSV.W) did not exhibit any bispecific binding activity, indicating that no background binding or a specific binding activity is produced.

10

\* All tested supernatants of COS-1 cells transfected with an expression vector encoding one of the  $VH_A$ - $VH_B$ - $VH_C$  fragments (pSV.K and pSV.M) and an expression vector encoding one of the  $VL_A$ - $VL_B$ - $VL_C$  fragments (pSV.V and pSV.W) showed significant levels of all three bifunctional binding activities LYSOX, GOSA and LYSAN.

15

These results show that COS cells transfected with expression vectors encoding  $VH_A$ - $VH_B$ - $VH_C$  and expression vectors encoding  $VL_A$ - $VL_B$ - $VL_C$  fragments produce and secrete molecules that contain three binding activities. In this example those three activities are: Glucose Oxidase binding, S.sanguis binding and Lysozyme binding. Furthermore, the results illustrated in Figure 10 clearly show that at least two of these binding activity are present in one self assembling molecular complex. In this example those combinations are: GOSA (Glucose Oxidase + S.sanguis), LYSOX (Lysozyme + Glucose Oxidase) and LYSAN (Lysozyme + S.sanguis).

20

25

30

**Example 3: Trifunctional binding activity of Golyzan triple heads**

35 This example describes experiments that show that the three types of bispecific binding activity that are produced by COS-1 cells transfected with expression plasmids encoding the corresponding  $VH_A$ - $VH_B$ - $VH_C$  and  $VL_A$ -

- 25 -

VL<sub>B</sub>-VL<sub>C</sub> genes fragments are present in one self assembling molecular complex.

5 Golysan.A (VHlys-VH4715-VH3418 + VLlys-VL4715-VL3418.2t)  
and Golysan.B (VHlys-VH4715-VH3418 + VL3418-VLlys-  
VL4715.2t) was purified by affinity chromatography.  
100ml supernatant of COS-1 cells transfected with  
expression plasmids pSV.M/pSV.V (Golysan.A) or  
pSV.M/pSV.W (Golysan.B) were loaded onto a Lysozyme-  
10 Sepharose column (CNBr-Sepharose, Pharmacia; column was  
prepared according to the manufacturer's instructions).  
After extensive washes with PBS the bound Golysan  
antibody fragments were eluted in 0.1M glycine buffer at  
pH=2.2. The fractions were neutralised with Tris and  
15 tested for the presence of trispecific binding activity.

As shown in Figure 11 no bispecific binding activity was  
detect in the column fall-through. All three bispecific  
binding activities (GOSA, LYSOX and LYSAN) were extracted  
20 from the COS-1 supernatant by passing over the Lysozyme  
affinity matrix. After acid elution all three bispecific  
binding activities (GOSA, LYSOX and LYSAN) were recovered  
from the column. Since both Golysan.A and B were  
affinity purified based on the ability to bind to  
25 Lysozyme, the finding that these molecules also bind  
S.sanguis and Glucose Oxidase shows that all three  
binding activities are present in one self assembling  
molecular complex.

### 30 Preparation 1.

#### Construction of the pGOSA.E double head expression vector

In the pGOSA expression vectors, the DNA fragments  
encoding both the V<sub>H</sub> and V<sub>L</sub> of the antibody are preceded  
35 by a ribosome binding site and a DNA sequence encoding  
the pelB signal sequence in an artificial dicistronic  
operon under the control of a single inducible promoter.

- 26 -

Expression of these constructs is driven by the inducible lacZ promoter. The nucleotide sequence of the HindIII-EcoRI inserts of the plasmids pUR.4124 (SEQ ID NO. 23), Fv.3418 (SEQ ID NO. 24), Fv.4715-myc (SEQ ID NO. 25) and scFv.4715-myc (SEQ ID NO. 26) constructs used for the generation of the bispecific antibody fragments are given in Figures 12-15, respectively. Moreover, a culture of *E. coli* cells harbouring plasmid scFv.4715-myc and a culture of *E. coli* cells harbouring plasmid Fv.3418 were deposited under the Budapest Treaty at the National Collection of Type Cultures (Central Public Health Laboratory) in London (United Kingdom) with deposition numbers NCTC 12916 and NCTC 12915, respectively.

In agreement with Rule 28 (4) EPC, or a similar arrangement for a State not being a Contracting State of the EPC, it is hereby requested that a sample of such deposit, when requested, will be submitted to an expert only.

The construction of pGOSA.E (see Figure 16 for the HindIII-EcoRI insert of pUC19) involved several cloning steps. The appropriate restriction sites in the various domains were introduced by PCR directed mutagenesis using the oligonucleotides listed in Table 1 below.

The construction of pGOSA.E involved several cloning steps that produced 4 intermediate constructs pGOSA.A to pGOSA.D (see Figures 17-21). The final expression vector pGOSA.E and the oligonucleotides in Table 1 have been designed to enable most specificities to be cloned into the final pGOSA.E construct (Figure 16c). The upstream  $V_H$  domain can be replaced by any PstI-BstEII  $V_H$  gene fragment obtained with oligonucleotides PCR.51 and PCR.89 (see Table 1). The oligonucleotides DBL.1 and DBL.2 (see Table 1) were designed to introduce SfiI and NheI restriction sites in the  $V_H$  gene fragments thus allowing

- 27 -

cloning of those  $V_H$  gene fragments into the *SfiI*-*NheI* sites as the downstream  $V_H$  domain. All  $V_L$  gene fragments obtained with oligonucleotides PCR.116 and PCR.90 (see Table 1) can be cloned into the position of the  $V_L$ .3418 gene fragment as a *SacI*-*XhoI* fragment. A complication here however is the presence of an internal *SacI* site in the  $V_H$ .3418 gene fragment. Oligonucleotides DBL.3 and DBL.9 (see Table 1) are designed to allow cloning of  $V_L$  gene fragments into the position of the  $V_L$ .4715 gene fragment as a *SalI*-*NotI* fragment.

#### **pGOSA.A**

This plasmid is derived from both the Fv.4715-myc construct (SEQ ID NO. 25) and the scFv.4715-myc construct (SEQ IN NO. 26). An *SfiI* restriction site was introduced between the DNA sequence encoding the (Gly,Ser)<sub>3</sub> linker and the gene fragment encoding the  $V_L$  of the scFv.4715-myc construct (see Figure 17). This was achieved by replacing the *BstEII*-*SacI* fragment of the latter construct by the fragment PCR-I *BstEII*/*SacI* (Figure 22) that contains an *SfiI* site between the DNA encoding the (Gly,Ser)<sub>3</sub> linker and the  $V_L$ .4715 gene fragment. The introduction of the *SfiI* site also introduced 4 additional amino acids (AlaGlySerAla) between the (Gly,Ser)<sub>3</sub> linker and  $V_L$ .4715 resulting in a (Gly,Ser)<sub>3</sub>AlaGlySerAla linker (linkerA). The oligonucleotides used to produce PCR-I (DBL.5 and DBL.7, see Table 1) were designed to match the sequence of the framework-3 region of  $V_H$ .4715 and to prime at the junction of the DNA encoding the (Gly,Ser)<sub>3</sub> linker and the  $V_L$ .4715 gene fragment, respectively. Thus pGOSA.A can be indicated as:

**pelB- $V_H$ 4715-linkerA-(*SfiI*)- $V_L$ 4715-myc.**

#### **pGOSA.B**

This plasmid is derived from plasmid Fv.3418 (see Figure 18). The *XhoI*-*EcoRI* fragment of plasmid Fv.3418

- 28 -

comprising the 3' end of DNA encoding framework-4 of the  $V_L$  including the stop codon was removed and replaced by the fragment PCR-IV *XhoI/EcoRI* (Figure 23). The oligonucleotides used to produce PCR-IV (DBL.8 and DBL.6, see Table 1) were designed to match the sequence at the junction of the  $V_L$  and the (Gly,Ser)<sub>3</sub> linker perfectly (DBL.8), and to be able to prime at the junction of the (Gly,Ser)<sub>3</sub> linker and the  $V_H$  in pUR.4124 (DBL.6). DBL.6 removed the *PstI* site in the  $V_H$  (silent mutation) and introduced a *SalI* restriction site at the junction of the (Gly,Ser)<sub>3</sub> linker and the  $V_H$ , thereby replacing the last Ser of the linker by a Val residue resulting in a (Gly,Ser)<sub>3</sub>Gly,Val linker (linkerV). Thus pGOSA.B can be indicated as:

15        **pelB- $V_H$ 3418 + pelB- $V_L$ 3418-linkerV-(*SalI-EcoRI*).**

#### **pGOSA.C**

This plasmid contains DNA encoding  $V_H$ .4715 linked by the (Gly,Ser)<sub>3</sub>AlaGlySerAla linker to  $V_H$ .3418 (see Figure 19), thus:

20                **pelB- $V_H$ 4715-linkerA- $V_H$ 3418.**

This construct was obtained by replacing the *SfiI-EcoRI* fragment from pGOSA.A encoding  $V_L$ .4715 by the fragment PCR-II *SfiI/EcoRI* containing the  $V_H$ .3418 gene. The oligonucleotides used to produce PCR-II (DBL.1 and DBL.2, see Table 1) hybridize in the framework-1 and framework-4 region of the gene encoding  $V_H$ .3418, respectively. DBL.1 was designed to remove the *PstI* restriction site (silent mutation) and to introduce an *SfiI* restriction site upstream of the  $V_H$  gene. DBL.2 destroys the *BstEII* restriction site in the framework-4 region and introduces an *NheI* restriction site downstream of the stopcodon.

#### **pGOSA.D**

35        This plasmid contains a dicistronic operon comprising the  $V_H$ .3418 gene and DNA encoding  $V_L$ .3418 linked by the

- 29 -

(Gly<sub>4</sub>Ser)<sub>2</sub>Gly<sub>4</sub>Val linker to V<sub>L</sub>.4715 (see Figure 20), thus:  
**pelB-V<sub>H</sub>.3418 + pelB-V<sub>L</sub>.3418-linkerV-V<sub>L</sub>.4715.**

5 This construct was obtained by digesting plasmid pGOSA.B  
with SalI-EcoRI and inserting the fragment PCR-V  
SalI/EcoRI (Figure 24) containing the V<sub>L</sub>.4715 gene. The  
oligonucleotides used to obtain PCR-V (DBL.3 and DBL.9,  
see Table 1) were designed to match the nucleotide  
10 the V<sub>L</sub>.4715 gene, respectively. DBL.3 removed the SacI  
site from the framework-1 region (silent mutation) and  
introduced a SalI restriction site upstream of the V<sub>L</sub>.4715  
gene. DBL.9 destroyed the XhoI restriction site in the  
framework-4 region of the V<sub>L</sub>.4715 gene (silent mutation)  
15 and introduced a NotI and an EcoRI restriction site  
downstream of the stop codon.

#### **pGOSA.E**

20 This plasmid contains a dicistronic operon comprising DNA  
encoding V<sub>H</sub>.4715 linked by the (Gly<sub>4</sub>Ser)<sub>2</sub>AlaGlySerAla  
linker to V<sub>H</sub>.3418 plus DNA encoding V<sub>L</sub>.3418 linked by the  
(Gly<sub>4</sub>Ser)<sub>2</sub>Gly<sub>4</sub>Val linker to V<sub>L</sub>.4715 (see Figure 21), thus:  
**pelB-V<sub>H</sub>.4715-linkerA-V<sub>H</sub>.3418 + pelB-V<sub>L</sub>.3418-linkerV-V<sub>L</sub>.4715.**

25 Both translational units are preceded by a ribosome  
binding site and DNA encoding a pelB leader sequence.  
This plasmid was obtained by a three-point ligation by  
mixing the vector resulting from pGOSA.D after removal of  
the V<sub>H</sub>.3418-encoding PstI-SacI insert with the PstI-NheI  
pGOSA.C insert containing V<sub>H</sub>.4715 linked to V<sub>H</sub>.3418 and  
30 the PCR-III NheI/SacI fragment (see Figure 25). The  
remaining PstI-SacI pGOSA.D vector contains the 5' end of  
the framework-1 region of V<sub>H</sub>.3418 upto the PstI  
restriction site and V<sub>L</sub>.3418 linked by the  
35 (Gly<sub>4</sub>Ser)<sub>2</sub>Gly<sub>4</sub>Val linker to V<sub>L</sub>.4715 starting from the SacI  
restriction site in V<sub>L</sub>.3418. The PstI-NheI pGOSA.C insert  
contains V<sub>H</sub>.4715 linked by the (Gly<sub>4</sub>Ser)<sub>2</sub>AlaGlySerAla

- 30 -

linker to V<sub>H</sub>.3418, starting from the *Pst*I restriction site in the framework-1 region in V<sub>H</sub>.4715. The *Nhe*I-*Sac*I PCR-III fragment provides the ribosome binding site and DNA encoding the *pelB* leader sequence for the V<sub>L</sub>.3418-  
5 (Gly,Ser)<sub>2</sub>Gly,Val-V<sub>L</sub>.4715 construct. The oligonucleotides DBL.10 and PCR.116 (see Table 1) used to generate PCR-III were designed to match the sequence upstream of the ribosome binding site of V<sub>L</sub>.4715 in Fv.4715 and to  
10 introduce an *Nhe*I restriction site (DBL.10), and to match the framework-4 region of V<sub>L</sub>.3418 (PCR.116).



- 31 -

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

- (A) NAME: UNILEVER PLC
- (B) STREET: Blackfriars
- (C) CITY: London
- (E) COUNTRY: United Kingdom
- (F) POSTAL CODE (ZIP): EC4P 4BQ
- (G) TELEPHONE: (01234) 222644
- (H) TELEFAX: (01234) 222633
- (I) TELEX: 82229 UNILAB G

(ii) TITLE OF INVENTION: Multivalent and multispecific antigen-binding protein

(iii) NUMBER OF SEQUENCES: 27

## (iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE:
  - PatentIn Release #1.0, Version #1.25 (EPO) (SEC ID NO. 1 to 18)
  - PatentIn Release #1.0, Version #1.30 (EPO) (SEC ID NO. 19 to 27)

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1745 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

AAGCTTGCAT GGAAATTCTA TTTCAAGGAG ACAGTCATAA TGAAATACCT ATTGCCTACG 60

GCAGCCGCTG GATTGTTATT ACTCGCTGCC CAACCAGCGA TGGCCCAGGT GCAGCTGCAG 120

- 32 -

GAGTCAGGGG GAGACTTAGT GAAGCCTGGA GGGTCCCTGA CACTCTCCTG TGCAACCTCT	180
GGATTCACTT TCAGTAGTTA TGCCTTTTCT TGGGTCCGCC AGACCTCAGA CAAGAGTCTG	240
GAGTGGGTCG CAACCATCAG TAGTACTGAT ACTTATACCT ATTATTCAGA CAATGTGAAG	300
GGGCGCTTCA CCATCTCCAG AGACAATGGC AAGAACACCC TGTACCTGCA AATGAGCAGT	360
CTGAAGTCTG AGGACACAGC CGTGTATTAC TGTGCAAGAC ATGGGTACTA TGGTAAAGGC	420
TATTTTGA CTACTGGGCCA AGGGACCACG GTCACCGTCT CCTCAGGTGG AGGCGGTTCA	480
GGCGGAGGTG GCTCTGGCGG TGGCGGATCG GCCGGTTCGG CCCAGGTCCA GCTGCAACAG	540
TCAGGACCTG AGCTGGTAAA GCCTGGGGCT TCAGTGAAGA TGTCTGCAA GGCTTCTGGA	600
TACACATTCA CTAGCTATGT TATGCACTGG GTGAAACAGA AGCCTGGGCA GGGCCTTGAG	660
TGGATTGGAT ATATTTATCC TTACAATGAT GGTACTAAGT ACAATGAGAA GTTCAAAGGC	720
AAGGCCACAC TGACTTCAGA CAAATCCTCC AGCACAGCCT ACATGGAGCT CAGCAGCCTG	780
ACCTCTGAGG ACTCTGCGGT CTATTACTGT TCAAGACGCT TTGACTACTG GGGCCAAGGG	840
ACCACCGTCA CCGTCTCCTC ATAATAAGCT AGCGGAGCTG CATGCAAATT CTATTTCAG	900
GAGACAGTCA TAATGAAATA CCTATTGCCT ACGGCAGCCG CTGGATTGTT ATTACTCGCT	960
GCCCAACCAG CGATGGCCGA CATCGAGCTC ACCCAGTCTC CATCTTCCAT GTATGCATCT	1020
CTAGGAGAGA GAATCACTAT CACTTGCAAG GCGAGTCAGG ACATTAATAC CTATTTAACC	1080
TGGTTCCAGC AGAAACCAGG GAAATCTCCC AAGACCCTGA TCTATCGTGC AAACAGATTG	1140
CTAGATGGGG TCCCATCAAG GTTCAGTGGC AGTGGATCTG GGCAAGATTA TTCTCTCACC	1200
ATCAGCAGCC TGGACTATGA AGATATGGGA ATTTATTATT GTCTACAATA TGATGAGTTG	1260
TACACGTTCTG GAGGGGGGAC CAAGCTCGAG ATCAAACGGG GTGGAGGCGG TTCAGGCGGA	1320
GGTGGCTCTG GCGGTGGCGG AGTCGACATC GAACTCACTC AGTCTCCATT CTCCCTGACT	1380
GTGACAGCAG GAGAGAAGGT CACTATGAAT TGCAAGTCCG GTCAGAGTCT GTTAAACAGT	1440

- 33 -

GTAAATCAGA GGAAC TACTT GACCTGGTAC CAGCAGAAGC CAGGGCAGCC TCCTAAACTG 1500  
TTGATCTACT GGGCATCCAC TAGGGAATCT GGAGTCCCTG ATCGCTTCAC AGCCAGTGG A 1560  
TCTGGAACAG ATTTCACTCT CACCATCAGC AGTGTGCAGG CTGAAGACCT GGCAGTTTAT 1620  
TACTGTCAGA ATGATTATAC TTATCCGTTT ACGTTCGGAG GGGGGACCAA GCTCGAAATC 1680  
AAACGGGGAT CCGGTAGCGG GAACTCCGGT AAGGGGTACC TGAAGTAATA AGCGGCCGCG 1740  
AATTC 1745

## (2) INFORMATION FOR SEQ ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 894 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

AAGCTTGCAT GCAAATTCTA TTCAAGGAG ACAGTCATAA TGAAATACCT ATTGCCTACG 60  
GCAGCCGCTG GATTGTTATT ACTCGCTGCC CAACCAGCGA TGGCCCAGGT GCAGCTGCAG 120  
GAGTCAGGAC CTGGCCTGGT GCGGCCCTCA CAGAGCCTGT CCATCACATG CACCGTCTCA 180  
GGGTTCTCAT TAACCGGCTA TGGTGTAAC TGGGTTGCC AGCCTCCAGG AAAGGGTCTG 240  
GAGTGGCTGG GAATGATTG GGGTGATGGA AACACAGACT ATAATTCAGC TCTCAAATCC 300  
AGACTGAGCA TCAGCAAGGA CAACTCCAAG AGCCAAGTTT TCTTAAAAAT GAACAGTCTG 360  
CACTGATG ACACAGCCAG GTACTACTGT GCCAGAGAGA GAGATTATAG GCTTGACTAC 420  
TGGGGCGAAG GCACCACGGT CACCGTCTCC TCAGGTGGAG GCGGTTGAG CGGAGGTGGC 480  
TCTGGCGGTG GCGGATCGGA CATCGAGCTC ACCCAGTCTC CAGCCTCCCT TTCTGCGTCT 540  
GTGGGAGAAA CTGTCACCAT CACATGTCGA GCAAGTGGGA ATATTCACAA TTATTTAGCA 600

- 34 -

TGGTATCAGC AGAAACAGGG AAAATCTCCT CAGCTCCTGG TCTATTATAC AACAACTTA 660  
GCAGATGGTG TGCCATCAAG GTTCAGTGGC AGTGGATCAG GAACACAATA TTCTCTCAAG 720  
ATCAACAGCC TGCAACCTGA AGATTTTGGG AGTTATTACT GTCAACATTT TTGGAGTACT 780  
CCTCGGACGT TCGGTGGAGG CACCAAGCTC GAGATCAAAC GGGAACAAAA ACTCATCTCA 840  
GAAGAGGATC TGAATTAATA AGATCAAACG GTAATAAGGA TCCAGCTCGA ATTC 894

## (2) INFORMATION FOR SEQ ID NO: 3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 930 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

AAGCTTGCAT GCAAATTCTA TTTCAAGGAG ACAGTCATAA TGAAATACCT ATTGCCTACG 60  
GCAGCCGCTG GATTGTTATT ACTCGCTGCC CAACCGGCCA TGGCCCAGGT GCAGCTGCAG 120  
GAGTCAGGGG GAGACTTAGT GAAGCCTGGA GGGTCCCTGA CACTCTCCTG TGCAACCTCT 180  
GGATTCACCT TCAGTAGTTA TGCCTTTTCT TGGGTCCGCC AGACCTCAGA CAAGAGTCTG 240  
GAGTGGGTCG CAACCATCAG TAGTACTGAT ACTTATACCT ATTATTCAGA CAATGTGAAG 300  
GGGCGCTTCA CCATCTCCAG AGACAATGGC AAGAACACCC TGTACCTGCA AATGAGCAGT 360  
CTGAAGTCTG AGGACACAGC CGTGTATTAC TGTGCAAGAC ATGGGTACTA TGGTAAAGGC 420  
TATTTTGACT ACTGGGGCCA AGGGACCACG GTCACCGTCT CTCAGGTGG AGGCGGTTCA 480  
GGCGGAGGTG GCTCTGGCGG TGGCGGATCG GACATCGAGC TCACTCAGTC TCCATTCTCC 540  
CTGACTGTGA CAGCAGGAGA GAAGGTCACT ATGAATTGCA AGTCCGGTCA GAGTCTGTTA 600  
AACAGTGTA ATCAGAGGAA CTACTTGACC TGGTACCAGC AGAAGCCAGG GCAGCCTCCT 660

- 35 -

AAACTGTTGA TCTACTGGGC ATCCACTAGG GAATCTGGAG TCCCTGATCG CTTCACAGCC 720  
AGTGGATCTG GAACAGATTT CACTCTCACC ATCAGCAGTG TGCAGGCTGA AGACCTGGCA 780  
GTTTATTACT GTCAGAATGA TTATACTTAT CCGTTCACGT TCGGAGGGGG GACCAAGCTC 840  
GAGATCAAAC GGGGATCCGG TAGCGGGAAC TCCGGTAAGG GGTACCTGAA GTAATAAGAT 900  
CAAACGGTAA TAAGGATCCA GCTCGAATTC 930

## (2) INFORMATION FOR SEQ ID NO: 4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 156 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CCATGGGATG GAGCTGTATC ATCCTCTTCT TGGTAGCAAC AGCTACAGGT AAGGGGCTCA 60  
CAGTAGCAGG CTTGAGGTCT GGACATATAT ATGGGTGACA ATGACATCCA CTTTGCCTTT 120  
CTCTCCACAG GTGTCCACTC CCAGGTCCAA CTGCAG 156

## (2) INFORMATION FOR SEQ ID NO: 5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AGGTSMAMCT GCAGSAGTCW GG 22

## (2) INFORMATION FOR SEQ ID NO: 6:

- 36 -

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

TGAGGAGACG GTGACCGTGG TCCCTTGGCC CC

32

## (2) INFORMATION FOR SEQ ID NO: 7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GACATTGAGC TCACCCAGTC TCCA

24

## (2) INFORMATION FOR SEQ ID NO: 8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GTTAGATCTC GAGCTTGGTC CC

22

## (2) INFORMATION FOR SEQ ID NO: 9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs

- 37 -

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CAGGATCCGG CCGGTTCCGC CCAGGTCCAG CTGCAACAGT CAGGA

45

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CTACATGAAT TCGCTAGCTT ATTATGAGGA GACGGTGACG GTGGTCCCTT GGC

53

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

ATTGGAGTCG ACATCGAACT CACTCAGTCT CCATTCTCC

39

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

- 38 -

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

CGAATTCGGA TCCCCGTTG ATTCGAGCT TGGTCC

36

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 45 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GAGCGCGAGC TCGCCGAAC CGGCCGATCC GCCACCGCCA GAGCC

45

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

ATTGTCGAAT TCGTCGACTC CGCCACCGCC AGAGCC

36

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 57 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear



- 39 -

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

AGCTTCTAGA CCACCATGGA AACTGCAGA GCTCAAAGC TAGCGCGGCG GCTCTAG

57

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 57 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

AATTCTAGAG CGGCCGCGCT AGCTTTTGAG CTCTGCAGTT TTCCATGGTG GTCTAGA

57

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

ACGGGTGAGC TCGATGTCGG AGTGGACACC TGTGGAGAGA

40

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- 40 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GGAAACAGCT ATGACCATGA TTAC

24

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:

(B) CLONE: primer DBL.7

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

CACCATCTCC AGAGACAATG GCAAG

25

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:

(B) CLONE: primer DBL.8

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

ACCAAGCTCG AGATCAAACG GGG

23

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 base pairs

- 41 -

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:

(B) CLONE: primer DBL.9

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

TGAAGTGAAT TCGCGGCCGC TTATTACCGT TTGATTTCGA GCTTGGTCCC

50

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(vii) IMMEDIATE SOURCE:

(B) CLONE: primer DBL.10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

TAATAAGCTA GCGGAGCTGC ATGCAAATTC TATTTTC

36

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 737 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "cDNA domains with synthetic linker(s)"

(vii) IMMEDIATE SOURCE:

- 42 -

(B) CLONE: EcoRI-HindIII insert of pUR4124

## (ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION:11..730

(D) OTHER INFORMATION:/product= "VLlys-GS-VHlys"

## (ix) FEATURE:

(A) NAME/KEY: mat\_peptide

(B) LOCATION:11..334

(D) OTHER INFORMATION:/product= "VLlys"

## (ix) FEATURE:

(A) NAME/KEY: misc\_RNA

(B) LOCATION:335..379

(D) OTHER INFORMATION:/product= "(Gly4Ser)3 linker"

## (ix) FEATURE:

(A) NAME/KEY: mat\_peptide

(B) LOCATION:380..727

(D) OTHER INFORMATION:/product= "VHlys"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

GAATTCGGCC GAC ATC GAG CTC ACC CAG TCT CCA GCC TCC CTT TCT GCG	49
Asp Ile Glu Leu Thr Gln Ser Pro Ala Ser Leu Ser Ala	
1 5 10	
TCT GTG GGA GAA ACT GTC ACC ATC ACA TGT CGA GCA AGT GGG AAT ATT	97
Ser Val Gly Glu Thr Val Thr Ile Thr Cys Arg Ala Ser Gly Asn Ile	
15 20 25	
CAC AAT TAT TTA GCA TGG TAT CAG CAG AAA CAG GGA AAA TCT CCT CAG	145
His Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Gln Gly Lys Ser Pro Gln	
30 35 40 45	
CTC CTG GTC TAT TAT ACA ACA ACC TTA GCA GAT GGT GTG CCA TCA AGG	193
Leu Leu Val Tyr Tyr Thr Thr Thr Leu Ala Asp Gly Val Pro Ser Arg	
50 55 60	
TTC AGT GGC AGT GGA TCA GGA ACA CAA TAT TCT CTC AAG ATC AAC AGC	241
Phe Ser Gly Ser Gly Ser Gly Thr Gln Tyr Ser Leu Lys Ile Asn Ser	
65 70 75	

- 43 -

CTG CAA CCT GAA GAT TTT GGG AGT TAT TAC TGT CAA CAT TTT TGG AGT Leu Gln Pro Glu Asp Phe Gly Ser Tyr Tyr Cys Gln His Phe Trp Ser 80 85 90	289
ACT CCT CGG ACG TTC GGT GGA GGG ACC AAG CTC GAG ATC AAA CGG GGT Thr Pro Arg Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg Gly 95 100 105	337
GGA GGC GGT TCA GGC GGA GGT GGC TCT GGC GGT GGC GGA TCG CAG GTG Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gln Val 110 115 120 125	385
CAG CTG CAG GAG TCA GGA CCT GGC CTG GTG GCG CCC TCA CAG AGC CTG Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Ala Pro Ser Gln Ser Leu 130 135 140	433
TCC ATC ACA TGC ACC GTC TCA GGG TTC TCA TTA ACC GGC TAT GGT GTA Ser Ile Thr Cys Thr Val Ser Gly Phe Ser Leu Thr Gly Tyr Gly Val 145 150 155	481
AAC TGG GTT CGC CAG CCT CCA GGA AAG GGT CTG GAG TGG CTG GGA ATG Asn Trp Val Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Leu Gly Met 160 165 170	529
ATT TGG GGT GAT GGA AAC ACA GAC TAT AAT TCA GCT CTC AAA TCC AGA Ile Trp Gly Asp Gly Asn Thr Asp Tyr Asn Ser Ala Leu Lys Ser Arg 175 180 185	577
CTG AGC ATC AGC AAG GAC AAC TCC AAG AGC CAA GTT TTC TTA AAA ATG Leu Ser Ile Ser Lys Asp Asn Ser Lys Ser Gln Val Phe Leu Lys Met 190 195 200 205	625
AAC AGT CTG CAC ACT GAT GAC ACA GCC AGG TAC TAC TGT GCC AGA GAG Asn Ser Leu His Thr Asp Asp Thr Ala Arg Tyr Tyr Cys Ala Arg Glu 210 215 220	673
AGA GAT TAT AGG CTT GAC TAC TGG GGC CAA GGG ACC ACG GTC ACC GTC Arg Asp Tyr Arg Leu Asp Tyr Trp Gly Gln Gly Thr Thr Val Thr Val 225 230 235	721
TCC TCA TGA TAAGCTT Ser Ser *	737

- 44 -

240

## (2) INFORMATION FOR SEQ ID NO: 24:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 920 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "cDNA domains with synthetic linker(s)"

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: HindIII-EcoRI insert Fv.3418

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION:36..443
- (D) OTHER INFORMATION:/product= "pelB-VH3418"

## (ix) FEATURE:

- (A) NAME/KEY: sig\_peptide
- (B) LOCATION:36..101
- (D) OTHER INFORMATION:/product= "pectate lyase"

## (ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) LOCATION:102..440
- (D) OTHER INFORMATION:/product= "VH3418"

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION:495..884
- (D) OTHER INFORMATION:/product= "pelB-VL4318"

## (ix) FEATURE:

- (A) NAME/KEY: sig\_peptide
- (B) LOCATION:495..560
- (D) OTHER INFORMATION:/product= "pectate lyase"

## (ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) LOCATION:561..881

- 45 -

(D) OTHER INFORMATION: /product= "VL3418"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

AAGCTTGCAA ATTCTATTTC AAGGAGACAG TCATA ATG AAA TAC CTA TTG CCT	53
Met Lys Tyr Leu Leu Pro	
-22 -20	
ACG GCA GCC GCT GGA TTG TTA TTA CTC GCT GCC CAA CCA GCG ATG GCC	101
Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro Ala Met Ala	
-15 -10 -5	
CAG GTG CAG CTG CAG CAG TCA GGA CCT GAG CTG GTA AAG CCT GGG GCT	149
Gln Val Gln Leu Gln Gln Ser Gly Pro Glu Leu Val Lys Pro Gly Ala	
1 5 10 15	
TCA GTG AAG ATG TCC TGC AAG GCT TCT GGA TAC ACA TTC ACT AGC TAT	197
Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr	
20 25 30	
GTT ATG CAC TGG GTG AAA CAG AAG CCT GGG CAG GGC CTT GAG TGG ATT	245
Val Met His Trp Val Lys Gln Lys Pro Gly Gln Gly Leu Glu Trp Ile	
35 40 45	
GGA TAT ATT TAT CCT TAC AAT GAT GGT ACT AAG TAC AAT GAG AAG TTC	293
Gly Tyr Ile Tyr Pro Tyr Asn Asp Gly Thr Lys Tyr Asn Glu Lys Phe	
50 55 60	
AAA GGC AAG GCC ACA CTG ACT TCA GAC AAA TCC TCC AGC ACA GCC TAC	341
Lys Gly Lys Ala Thr Leu Thr Ser Asp Lys Ser Ser Ser Thr Ala Tyr	
65 70 75 80	
ATG GAG CTC AGC AGC CTG ACC TCT GAG GAC TCT GCG GTC TAT TAC TGT	389
Met Glu Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys	
85 90 95	
TCA AGA CGC TTT GAC TAC TGG GGC CAA GGG ACC ACG GTC ACC GTC TCC	437
Ser Arg Arg Phe Asp Tyr Trp Gly Gln Gly Thr Thr Val Thr Val Ser	
100 105 110	
TCA TAA TAAGAGCTAT GGGAGCTTGC ATGCAAATTC TATTTCAAGG AGACAGTCAT	493
Ser *	
A ATG AAA TAC CTA TTG CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC	539

- 46 -

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu  
-22        -20                        -15                        -10

GCT GCC CAA CCA GCG ATG GCC GAC ATC GAG CTC ACC CAG TCT CCA TCT 587  
Ala Ala Gln Pro Ala Met Ala Asp Ile Glu Leu Thr Gln Ser Pro Ser  
-5 1 5

TCC ATG TAT GCA TCT CTA GGA GAG AGA ATC ACT ATC ACT TGC AAG GCG 635  
Ser Met Tyr Ala Ser Leu Gly Glu Arg Ile Thr Ile Thr Cys Lys Ala  
10 15 20 25

AGT CAG GAC ATT AAT ACC TAT TTA ACC TGG TTC CAG CAG AAA CCA GGG 683  
Ser Gln Asp Ile Asn Thr Tyr Leu Thr Trp Phe Gln Gln Lys Pro Gly  
30 35 40

AAA TCT CCC AAG ACC CTG ATC TAT CGT GCA AAC AGA TTG CTA GAT GGG 731  
Lys Ser Pro Lys Thr Leu Ile Tyr Arg Ala Asn Arg Leu Leu Asp Gly  
45 50 55

GTC CCA TCA AGG TTC AGT GGC AGT GGA TCT GGG CAA GAT TAT TCT CTC 779  
Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Gln Asp Tyr Ser Leu  
60 65 70

ACC ATC AGC AGC CTG GAC TAT GAA GAT ATG GGA ATT TAT TAT TGT CTA 827  
Thr Ile Ser Ser Leu Asp Tyr Glu Asp Met Gly Ile Tyr Tyr Cys Leu  
75 80 85

CAA TAT GAT GAG TTG TAC ACG TTC GGA GGG GGG ACC AAG CTC GAG ATC 875  
Gln Tyr Asp Glu Leu Tyr Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile  
90 95 100 105

AAA CGG TAA TAATGATCAA ACGGTATAAG GATCCAGCTC GAATTC 920  
Lys Arg \*

(2) INFORMATION FOR SEQ ID NO: 25:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 999 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "cDNA domains with synthetic linker(s)"



- 47 -

## (vii) IMMEDIATE SOURCE:

(B) CLONE: HindIII-EcoRI insert of Fv.4715-myc

## (ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION:40..468

(D) OTHER INFORMATION:/product= "pelB-VH4715"

## (ix) FEATURE:

(A) NAME/KEY: sig\_peptide

(B) LOCATION:40..105

(D) OTHER INFORMATION:/product= "pectate lyase"

## (ix) FEATURE:

(A) NAME/KEY: mat\_peptide

(B) LOCATION:106..465

(D) OTHER INFORMATION:/product= "VH4715"

## (ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION:520..963

(D) OTHER INFORMATION:/product= "pelB-VL4715-myc"

## (ix) FEATURE:

(A) NAME/KEY: sig\_peptide

(B) LOCATION:520..585

(D) OTHER INFORMATION:/product= "pectate lyase"

## (ix) FEATURE:

(A) NAME/KEY: mat\_peptide

(B) LOCATION:586..927

(D) OTHER INFORMATION:/product= "VL4715"

## (ix) FEATURE:

(A) NAME/KEY: misc\_RNA

(B) LOCATION:928..960

(D) OTHER INFORMATION:/product= "myc-tag"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

AAGCTTGCAT GCAAATTCTA TTTCAAGGAG ACAGTCATA ATG AAA TAC CTA TTG  
Met Lys Tyr Leu Leu

54

-22 -20

- 48 -

CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC GCT GCC CAA CCA GCG ATG	102
Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro Ala Met	
-15 -10 -5	
GCC CAG GTG CAG CTG CAG GAG TCA GGG GGA GAC TTA GTG AAG CCT GGA	150
Ala Gln Val Gln Leu Gln Glu Ser Gly Gly Asp Leu Val Lys Pro Gly	
1 5 10 15	
GGG TCC CTG ACA CTC TCC TGT GCA ACC TCT GGA TTC ACT TTC AGT AGT	198
Gly Ser Leu Thr Leu Ser Cys Ala Thr Ser Gly Phe Thr Phe Ser Ser	
20 25 30	
TAT GCC TTT TCT TGG GTC CGC CAG ACC TCA GAC AAG AGT CTG GAG TGG	246
Tyr Ala Phe Ser Trp Val Arg Gln Thr Ser Asp Lys Ser Leu Glu Trp	
35 40 45	
GTC GCA ACC ATC AGT AGT ACT GAT ACT TAT ACC TAT TAT TCA GAC AAT	294
Val Ala Thr Ile Ser Ser Thr Asp Thr Tyr Thr Tyr Tyr Ser Asp Asn	
50 55 60	
GTG AAG GGG CGC TTC ACC ATC TCC AGA GAC AAT GGC AAG AAC ACC CTG	342
Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Gly Lys Asn Thr Leu	
65 70 75	
TAC CTG CAA ATG AGC AGT CTG AAG TCT GAG GAC ACA GCC GTG TAT TAC	390
Tyr Leu Gln Met Ser Ser Leu Lys Ser Glu Asp Thr Ala Val Tyr Tyr	
80 85 90 95	
TGT GCA AGA CAT GGG TAC TAT GGT AAA GGC TAT TTT GAC TAC TGG GGC	438
Cys Ala Arg His Gly Tyr Tyr Gly Lys Gly Tyr Phe Asp Tyr Trp Gly	
100 105 110	
CAA GGG ACC ACG GTC ACC GTC TCC TCA TAA TAAGAGCTAT GGGAGCTTGC	488
Gln Gly Thr Thr Val Thr Val Ser Ser *	
115 120	
ATGCAAATTC TATTTCAAGG AGACAGTCAT A ATG AAA TAC CTA TTG CCT ACG	540
Met Lys Tyr Leu Leu Pro Thr	
-22 -20	
GCA GCC GCT GGA TTG TTA TTA CTC GCT GCC CAA CCA GCG ATG GCC GAC	588
Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro Ala Met Ala Asp	
-15 -10 -5 1	

- 49 -

ATC GAG CTC ACT CAG TCT CCA TTC TCC CTG ACT GTG ACA GCA GGA GAG Ile Glu Leu Thr Gln Ser Pro Phe Ser Leu Thr Val Thr Ala Gly Glu	636
5 10 15	
AAG GTC ACT ATG AAT TGC AAG TCC GGT CAG AGT CTG TTA AAC AGT GTA Lys Val Thr Met Asn Cys Lys Ser Gly Gln Ser Leu Leu Asn Ser Val	684
20 25 30	
AAT CAG AGG AAC TAC TTG ACC TGG TAC CAG CAG AAG CCA GGG CAG CCT Asn Gln Arg Asn Tyr Leu Thr Trp Tyr Gln Gln Lys Pro Gly Gln Pro	732
35 40 45	
CCT AAA CTG TTG ATC TAC TGG GCA TCC ACT AGG GAA TCT GGA GTC CCT Pro Lys Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val Pro	780
50 55 60 65	
GAT CGC TTC ACA GCC AGT GGA TCT GGA ACA GAT TTC ACT CTC ACC ATC Asp Arg Phe Thr Ala Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile	828
70 75 80	
AGC AGT GTG CAG GCT GAA GAC CTG GCA GTT TAT TAC TGT CAG AAT GAT Ser Ser Val Gln Ala Glu Asp Leu Ala Val Tyr Tyr Cys Gln Asn Asp	876
85 90 95	
TAT ACT TAT CCG TTC ACG TTC GGA GGG GGG ACC AAG CTC GAG ATC AAA Tyr Thr Tyr Pro Phe Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys	924
100 105 110	
CGG GAA CAA AAA CTC ATC TCA GAA GAG GAT CTG AAT TAA TAAGATCAAA Arg Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn *	973
115 120 125	
CGGTAATAAG GATCCAGCTC GAATTC	999

## (2) INFORMATION FOR SEQ ID NO: 26:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 924 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "cDNA domains with synthetic linker(s)"

- 50 -

(vii) IMMEDIATE SOURCE:

(B) CLONE: HindIII-EcoRI insert of scFv.4715-myc

(ix) **FEATURE:**

(A) NAME/KEY: sig\_peptide

(B) LOCATION:40..105

(D) OTHER INFORMATION: /product= "pectate lyase"

**(ix) FEATURE:**

(A) NAME/KBY: mat\_peptide

(B) LOCATION: 106..465

(D) OTHER INFORMATION: /product= "VH4715"

**(ix) FEATURE:**

(A) NAME/KEY: misc RNA

(B) LOCATION: 466..510

(D) OTHER INFORMATION: /product= "(Gly4Ser)3-linker"

**(ix) FEATURE:**

(A) NAME/KEY: mat\_peptide

(B) LOCATION: 511..852

(D) OTHER INFORMATION: /product= "VL4715"

**(ix) FEATURE:**

(A) NAME/KEY: misc\_RNA

(B) LOCATION: 853..885

(D) OTHER INFORMATION: /product= "myc-tag"

**(ix) FEATURE:**

(A) NAME/KEY: CDS

- (B) LOCATION:40..888

(D) OTHER INFORMATION: /product= "pelB-VH4715- (Gly4Ser) 3-VL4715-myc"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

AAGCTTG CAT GCAAATTCTA TTTCAAGGAG ACAGTCATA ATG AAA TAC CTA TTG 54  
 Met Lys Tyr Leu Leu  
 -22 -20

CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC GCT GCC CAA CCA GCG ATG 102  
 Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro Ala Met  
 -15 -10 -5

- 51 -

GCC CAG GTG CAG CTG CAG GAG TCA GGG GGA GAC TTA GTG AAG CCT GGA	150
Ala Gln Val Gln Leu Gln Glu Ser Gly Gly Asp Leu Val Lys Pro Gly	
1 5 10 15	
GGG TCC CTG ACA CTC TCC TGT GCA ACC TCT GGA TTC ACT TTC AGT AGT	198
Gly Ser Leu Thr Leu Ser Cys Ala Thr Ser Gly Phe Thr Phe Ser Ser	
20 25 30	
TAT GCC TTT TCT TGG GTC CGC CAG ACC TCA GAC AAG AGT CTG GAG TGG	246
Tyr Ala Phe Ser Trp Val Arg Gln Thr Ser Asp Lys Ser Leu Glu Trp	
35 40 45	
GTC GCA ACC ATC AGT AGT ACT GAT ACT TAT ACC TAT TAT TCA GAC AAT	294
Val Ala Thr Ile Ser Ser Thr Asp Thr Tyr Thr Tyr Tyr Ser Asp Asn	
50 55 60	
GTG AAG GGG CGC TTC ACC ATC TCC AGA GAC AAT GGC AAG AAC ACC CTG	342
Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Gly Lys Asn Thr Leu	
65 70 75	
TAC CTG CAA ATG AGC AGT CTG AAG TCT GAG GAC ACA GCC GTG TAT TAC	390
Tyr Leu Gln Met Ser Ser Leu Lys Ser Glu Asp Thr Ala Val Tyr Tyr	
80 85 90 95	
TGT GCA AGA CAT GGG TAC TAT GGT AAA GGC TAT TTT GAC TAC TGG GGC	438
Cys Ala Arg His Gly Tyr Tyr Gly Lys Gly Tyr Phe Asp Tyr Trp Gly	
100 105 110	
CAA GGG ACC ACG GTC ACC GTC TCC TCA GGT GGA GGC GGT TCA GGC GGA	486
Gln Gly Thr Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly	
115 120 125	
GGT GGC TCT GGC GGT GGC GGA TCG GAC ATC GAG CTC ACT CAG TCT CCA	534
Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Glu Leu Thr Gln Ser Pro	
130 135 140	
TTC TCC CTG ACT GTG ACA GCA GGA GAG AAG GTC ACT ATG AAT TGC AAG	582
Phe Ser Leu Thr Val Thr Ala Gly Glu Lys Val Thr Met Asn Cys Lys	
145 150 155	
TCC GGT CAG AGT CTG TTA AAC AGT GTA AAT CAG AGG AAC TAC TTG ACC	630
Ser Gly Gln Ser Leu Leu Asn Ser Val Asn Gln Arg Asn Tyr Leu Thr	
160 165 170 175	

- 52 -

TGG TAC CAG CAG AAG CCA GGG CAG CCT CCT AAA CTG TTG ATC TAC TGG	678
Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Tyr Trp	
180 185 190	
GCA TCC ACT AGG GAA TCT GGA GTC CCT GAT CGC TTC ACA GCC AGT GGA	726
Ala Ser Thr Arg Glu Ser Gly Val Pro Asp Arg Phe Thr Ala Ser Gly	
195 200 205	
TCT GGA ACA GAT TTC ACT CTC ACC ATC AGC AGT GTG CAG GCT GAA GAC	774
Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Val Gln Ala Glu Asp	
210 215 220	
CTG GCA GTT TAT TAC TGT CAG AAT GAT TAT ACT TAT CCG TTC ACG TTC	822
Leu Ala Val Tyr Tyr Cys Gln Asn Asp Tyr Thr Tyr Pro Phe Thr Phe	
225 230 235	
GGA GGG GGG ACC AAG CTC GAG ATC AAA CGG GAA CAA AAA CTC ATC TCA	870
Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg Glu Gln Lys Leu Ile Ser	
240 245 250 255	
GAA GAG GAT CTG AAT TAA TAAGATCAAA CGGTAATAAG GATCCAGCTC GAATTC	924
Glu Glu Asp Leu Asn *	
260	

## (2) INFORMATION FOR SEQ ID NO: 27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1706 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "cDNA domains with synthetic linker(s)"

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: HindIII-EcoRI insert of pGOSA.B

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 40..864
- (D) OTHER INFORMATION: /product= "pelB-VH4715-LiA-VH3418"

## (ix) FEATURE:

- 53 -

- (A) NAME/KEY: sig\_peptide
- (B) LOCATION:40..105
- (D) OTHER INFORMATION:/product= "pectate lyase"

## (ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) LOCATION:106..465
- (D) OTHER INFORMATION:/product= "VH4715"

## (ix) FEATURE:

- (A) NAME/KEY: misc\_RNA
- (B) LOCATION:466..522
- (D) OTHER INFORMATION:/product= "linkerA  
(Gly4Ser)3AlaGlySerAla"

## (ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) LOCATION:523..861
- (D) OTHER INFORMATION:/product= "VH3418"

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION:913..1689
- (D) OTHER INFORMATION:/product= "pelB-VL3418-LiV-VL4715"

## (ix) FEATURE:

- (A) NAME/KEY: sig\_peptide
- (B) LOCATION:913..978
- (D) OTHER INFORMATION:/product= "pectate lyase"

## (ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) LOCATION:979..1299
- (D) OTHER INFORMATION:/product= "VL3418"

## (ix) FEATURE:

- (A) NAME/KEY: misc\_RNA
- (B) LOCATION:1300..1344
- (D) OTHER INFORMATION:/product= "linker V  
(Gly4Ser)2Gly4Val"

## (ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) LOCATION:1345..1686

- 54 -

(D) OTHER INFORMATION: /product= "VL4715"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

AAGCTTGCAT GGAAATTCTA TTTCAAGGAG ACAGTCATA ATG AAA TAC CTA TTG	54
Met Lys Tyr Leu Leu	
-22 -20	
CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC GCT GCC CAA CCA GCG ATG	102
Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro Ala Met	
-15 -10 -5	
GCC CAG GTG CAG CTG CAG GAG TCA GGG GGA GAC TTA GTG AAG CCT GGA	150
Ala Gln Val Gln Leu Gln Glu Ser Gly Gly Asp Leu Val Lys Pro Gly	
1 5 10 15	
GGG TCC CTG ACA CTC TCC TGT GCA ACC TCT GGA TTC ACT TTC AGT AGT	198
Gly Ser Leu Thr Leu Ser Cys Ala Thr Ser Gly Phe Thr Phe Ser Ser	
20 25 30	
TAT GCC TTT TCT TGG GTC CGC CAG ACC TCA GAC AAG AGT CTG GAG TGG	246
Tyr Ala Phe Ser Trp Val Arg Gln Thr Ser Asp Lys Ser Leu Glu Trp	
35 40 45	
GTC GCA ACC ATC AGT AGT ACT GAT ACT TAT ACC TAT TAT TCA GAC AAT	294
Val Ala Thr Ile Ser Ser Thr Asp Thr Tyr Thr Tyr Tyr Ser Asp Asn	
50 55 60	
GTG AAG GGG CGC TTC ACC ATC TCC AGA GAC AAT GGC AAG AAC ACC CTG	342
Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Gly Lys Asn Thr Leu	
65 70 75	
TAC CTG CAA ATG AGC AGT CTG AAG TCT GAG GAC ACA GCC GTG TAT TAC	390
Tyr Leu Gln Met Ser Ser Leu Lys Ser Glu Asp Thr Ala Val Tyr Tyr	
80 85 90 95	
TGT GCA AGA CAT GGG TAC TAT GGT AAA GGC TAT TTT GAC TAC TGG GGC	438
Cys Ala Arg His Gly Tyr Tyr Gly Lys Gly Tyr Phe Asp Tyr Trp Gly	
100 105 110	
CAA GGG ACC ACG GTC ACC GTC TCC TCA GGT GGA GGC GGT TCA GGC GGA	486
Gln Gly Thr Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly	
115 120 125	



- 55 -

GGT GGC TCT GGC GGT GGC GGA TCG GCC GGT TCG GCC CAG GTC CAG CTG Gly Gly Ser Gly Gly Gly Gly Ser Ala Gly Ser Ala Gln Val Gln Leu 130 135 140	534
CAA CAG TCA GGA CCT GAG CTG GTA AAG CCT GGG GCT TCA GTG AAG ATG Gln Gln Ser Gly Pro Glu Leu Val Lys Pro Gly Ala Ser Val Lys Met 145 150 155	582
TCC TGC AAG GCT TCT GGA TAC ACA TTC ACT AGC TAT GTT ATG CAC TGG Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr Val Met His Trp 160 165 170 175	630
GTG AAA CAG AAG CCT GGG CAG GGC CTT GAG TGG ATT GGA TAT ATT TAT Val Lys Gln Lys Pro Gly Gln Gly Leu Glu Trp Ile Gly Tyr Ile Tyr 180 185 190	678
CCT TAC AAT GAT GGT ACT AAG TAC AAT GAG AAG TTC AAA GGC AAG GCC Pro Tyr Asn Asp Gly Thr Lys Tyr Asn Glu Lys Phe Lys Gly Lys Ala 195 200 205	726
ACA CTG ACT TCA GAC AAA TCC TCC AGC ACA GCC TAC ATG GAG CTC AGC Thr Leu Thr Ser Asp Lys Ser Ser Ser Thr Ala Tyr Met Glu Leu Ser 210 215 220	774
AGC CTG ACC TCT GAG GAC TCT GCG GTC TAT TAC TGT TCA AGA CGC TTT Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys Ser Arg Arg Phe 225 230 235	822
GAC TAC TGG GGC CAA GGG ACC ACC GTC ACC GTC TCC TCA TAA Asp Tyr Trp Gly Gln Gly Thr Thr Val Thr Val Ser Ser * 240 245 250	864
TAAGCTAGCG GAGCTGCATG CAAATTCTAT TTCAAGGAGA CAGTCATA ATG AAA TAC Met Lys Tyr -22 -20	921
CTA TTG CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC GCT GCC CAA CCA Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro -15 -10 -5	969
GCG ATG GCC GAC ATC GAG CTC ACC CAG TCT CCA TCT TCC ATG TAT GCA Ala Met Ala Asp Ile Glu Leu Thr Gln Ser Pro Ser Ser Met Tyr Ala 1 5 10	1017

- 56 -

TCT CTA GGA GAG AGA ATC ACT ATC ACT TGC AAG GCG AGT CAG GAC ATT	1065
Ser Leu Gly Glu Arg Ile Thr Ile Thr Cys Lys Ala Ser Gln Asp Ile	
15 20 25	
AAT ACC TAT TTA ACC TGG TTC CAG CAG AAA CCA GGG AAA TCT CCC AAG	1113
Asn Thr Tyr Leu Thr Trp Phe Gln Gln Lys Pro Gly Lys Ser Pro Lys	
30 35 40 45	
ACC CTG ATC TAT CGT GCA AAC AGA TTG CTA GAT GGG GTC CCA TCA AGG	1161
Thr Leu Ile Tyr Arg Ala Asn Arg Leu Leu Asp Gly Val Pro Ser Arg	
50 55 60	
TTC AGT GGC AGT GGA TCT GGG CAA GAT TAT TCT CTC ACC ATC AGC AGC	1209
Phe Ser Gly Ser Gly Ser Gly Gln Asp Tyr Ser Leu Thr Ile Ser Ser	
65 70 75	
CTG GAC TAT GAA GAT ATG GGA ATT TAT TAT TGT CTA CAA TAT GAT GAG	1257
Leu Asp Tyr Glu Asp Met Gly Ile Tyr Tyr Cys Leu Gln Tyr Asp Glu	
80 85 90	
TTG TAC ACG TTC GGA GGG GGG ACC AAG CTC GAG ATC AAA CGG GGT GGA	1305
Leu Tyr Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg Gly Gly	
95 100 105	
GGC GGT TCA GGC GGA GGT GGC TCT GGC GGT GGC GGA GTC GAC ATC GAA	1353
Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Val Asp Ile Glu	
110 115 120 125	
CTC ACT CAG TCT CCA TTC TCC CTG ACT GTG ACA GCA GGA GAG AAG GTC	1401
Leu Thr Gln Ser Pro Phe Ser Leu Thr Val Thr Ala Gly Glu Lys Val	
130 135 140	
ACT ATG AAT TGC AAG TCC GGT CAG AGT CTG TTA AAC AGT GTA AAT CAG	1449
Thr Met Asn Cys Lys Ser Gly Gln Ser Leu Leu Asn Ser Val Asn Gln	
145 150 155	
AGG AAC TAC TTG ACC TGG TAC CAG CAG AAG CCA GGG CAG CCT CCT AAA	1497
Arg Asn Tyr Leu Thr Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys	
160 165 170	
CTG TTG ATC TAC TGG GCA TCC ACT AGG GAA TCT GGA GTC CCT GAT CGC	1545
Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val Pro Asp Arg	
175 180 185	

- 57 -

TTC ACA GCC AGT GGA TCT GGA ACA GAT TTC ACT CTC ACC ATC AGC AGT 1593  
Phe Thr Ala Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser  
190 195 200 205

GTG CAG GCT GAA GAC CTG GCA GTT TAT TAC TGT CAG AAT GAT TAT ACT 1641  
Val Gln Ala Glu Asp Leu Ala Val Tyr Tyr Cys Gln Asn Asp Tyr Thr  
210 215 220

TAT CCG TTC ACG TTC GGA GGG GGG ACC AAG CTC GAA ATC AAA CGG TAA 1689  
Tyr Pro Phe Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg \*  
225 230 235

TAAGCGGCCG CGAATTC 1706

**CLAIMS**

1. A multivalent antigen binding protein comprising:  
a first polypeptide comprising, in series,  
5 three or more variable domains of an antibody heavy chain; and  
a second polypeptide comprising, in series,  
three or more variable domains of an antibody light chain,  
10 said first and second polypeptides being linked by association of the respective heavy chain and light chain variable domains, each associated variable domain pair forming an antigen binding site.
- 15 2. A protein according to Claim 1 comprising a trivalent antigen binding protein.
- 20 3. A protein according to Claim 1 or Claim 2 wherein the variable domains of the antibody heavy chain of said first polypeptide are linked by a peptide linker and the variable domains of the antibody light chain of said second polypeptide are linked by a peptide linker.
- 25 4. A protein according to any one of Claims 1 to 3 wherein the associated variable domain pair binding sites are able to bind different epitopes from each other.
- 30 5. A protein according to any one of Claims 1 to 3 wherein the associated variable domain pair binding sites are able to bind the same epitope as each other.
- 35 6. Nucleotide sequences coding for the polypeptides of the multivalent antigen binding protein of any one

- 59 -

of the preceding claims.

7. Nucleotide sequences according to Claim 6 contained in one or more expression vectors.

5 8. A host cell transformed with a vector according to Claim 7, and capable of expression of the nucleotide sequences to produce the polypeptides of the multivalent antigen binding protein.

10 9. A host cell according to Claim 8 wherein the polypeptides on expression associate to form the multivalent antigen binding protein.

15 10. A process for preparing a multivalent antigen binding protein according to any one of Claims 1 to 5 comprising

20 (i) transforming one or more hosts by incorporating genes encoding said first and second polypeptides;  
(ii) expressing said genes and said host or hosts; and  
(iii) allowing said first and second polypeptides to associate to form the protein.

25 11. A protein according to any one of Claims 1 to 5 for use in medicine.

30 12. A diagnostic or therapeutic composition comprising a protein according to any one of Claims 1 to 5.

13. Use of composition according to Claim 12 in the preparation of an agent for use in diagnosis or therapy.

35 14. A method of diagnosis or therapy comprising administering a protein according to any one of

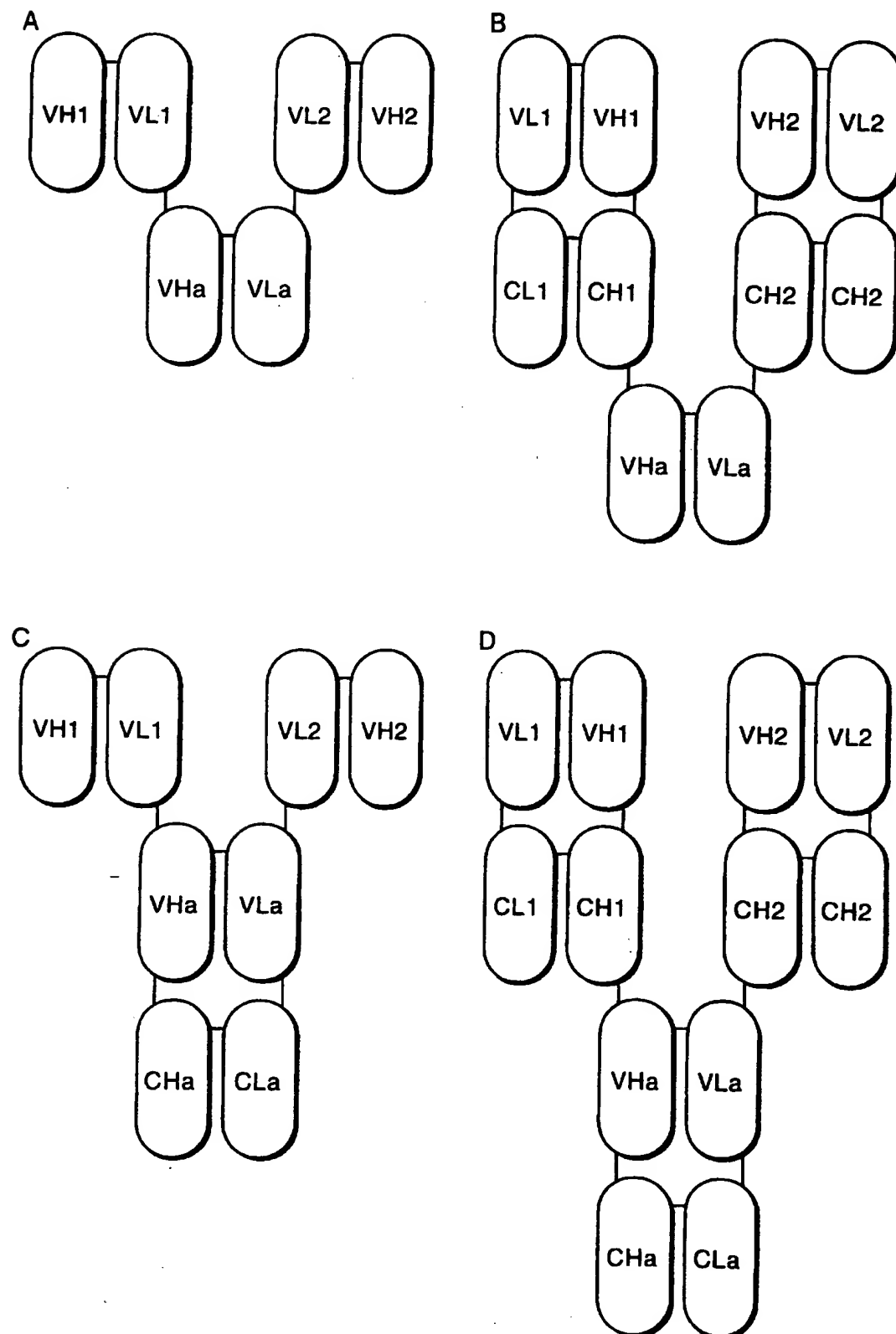
- 60 -

Claims 1 to 5.

15. Use of a protein according to any one of Claims 1 to 5 in an immunoassay method or for purification.

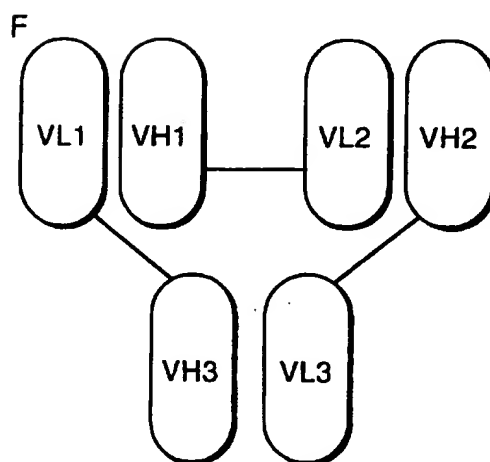
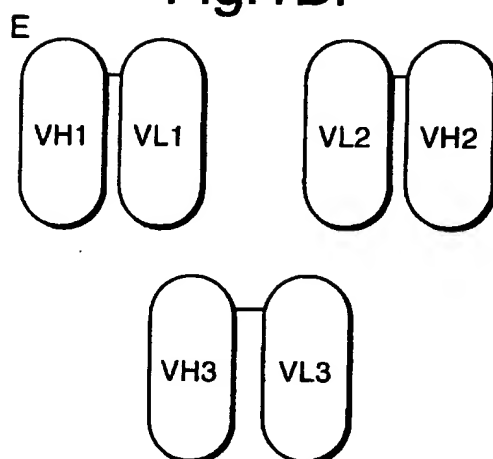
1/35

Fig.1A.



2/35

Fig. 1B.





3/35

## Fig.2A.

M K Y L L P T

**pelB leader** AAGCTTGCATGCAAATTCATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACG

A A A G L L L L A A Q P A M A Q V Q L Q

GCAGCCGCTGGATTGTTATTACTCGCTGCCCAACCAGCGATGGCCCAGGTGCAGCTGCAG

E S G G D L V K P G G S L T L S C A T S

GAGTCAGGGGGAGACTTAGTGAAGCCTGGAGGGTCCCTGACACTCTCCTGTGCAACCTCT

G F T F S S Y A F S W V R Q T S D K S L

GGATTCACTTTTCAGTAGTTATGCCTTTTCTTGGGTCCGCCAGACCTCAGACAAGAGTCTG

E W V A T L S S T D T Y T Y Y S D N V K

GAGTGGGTCGCAACCATCAGTAGTACTGATACTTATACCTATTATTTCAGACAATGTGAAG

**VH4715** G R F T L S E D N G R N T L Y L Q M E S

GGGCGCTTCACCATCTCCAGAGACAATGCAAGAACACCCTGTACCTGCAAATGAGCAGT

L K S E D T A V Y Y C A R H G Y Y G K E

CTGAAGTCTGAGGACACAGCCGTGTATTACTGTGCAAGACATGGGTACTATGGTAAAGGC

Y F D Y W G Q G T T V T V S S G G G G S

TATTTTGACTACTGGGGCCAAGGGACCACGGTCACCGTCTCCTCAGGTGGAGGCGGTTCA

**Linker** G G G G S G G G G S A G S A Q V Q L Q Q

GGCGGAGGTGGCTCTGGCGGTGGCGGATCGGCCGGTTCGGCCCAGGTCCAGCTGCAACAG

S G P E L V K P G A S V K M S C K A S G

TCAGGACCTGAGCTGGTAAAGCCTGGGGCTTCAGTGAAGATGTCTGCAAGGCTTCTGGA

Y T F T S Y V M H W V K Q K P G Q G L E

TACACATTCAGTATGTTATGCACTGGGTGAAACAGAAGCCTGGGCAGGCGCTTGAG

W I G Y T L P Y N D G T R Y N E K F K G

**VH3418** TGGATTGGATATATTTATCCTTACAATGATGGTACTAAGTACAATGAGAAGTCAAAGGC

K A T L T S D K S S S T A Y H E L S S L

AAGCCACACTGACTTCAGACAAATCCTCCAGCACAGCCTACATGGAGCTCAGCAGCCTG

T S E D S A V Y Y C S R R F D Y W G Q G

ACCTCTGAGGACTCTGCGGTCTATTACTGTTCAAGACGCTTTGACTACTGGGGCCAAGGG

T T V T V S S

ACCACCGTCACCGTCTCCTCATAATAAGCTAGCGGAGCTGCATGCAAATTCATTTCAAG

**pelB leader** M K Y L L P T A A A G L L L L A

GAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCT

A Q P A M A D I E L T Q S P S P M Y A S

GCCCAACCAGCGATGGCCGACATCGAGCTACCCAGTCTCCATCTTCCATGTATGCATCT

L G E R I T L T C K A S Q D T N T Y L T

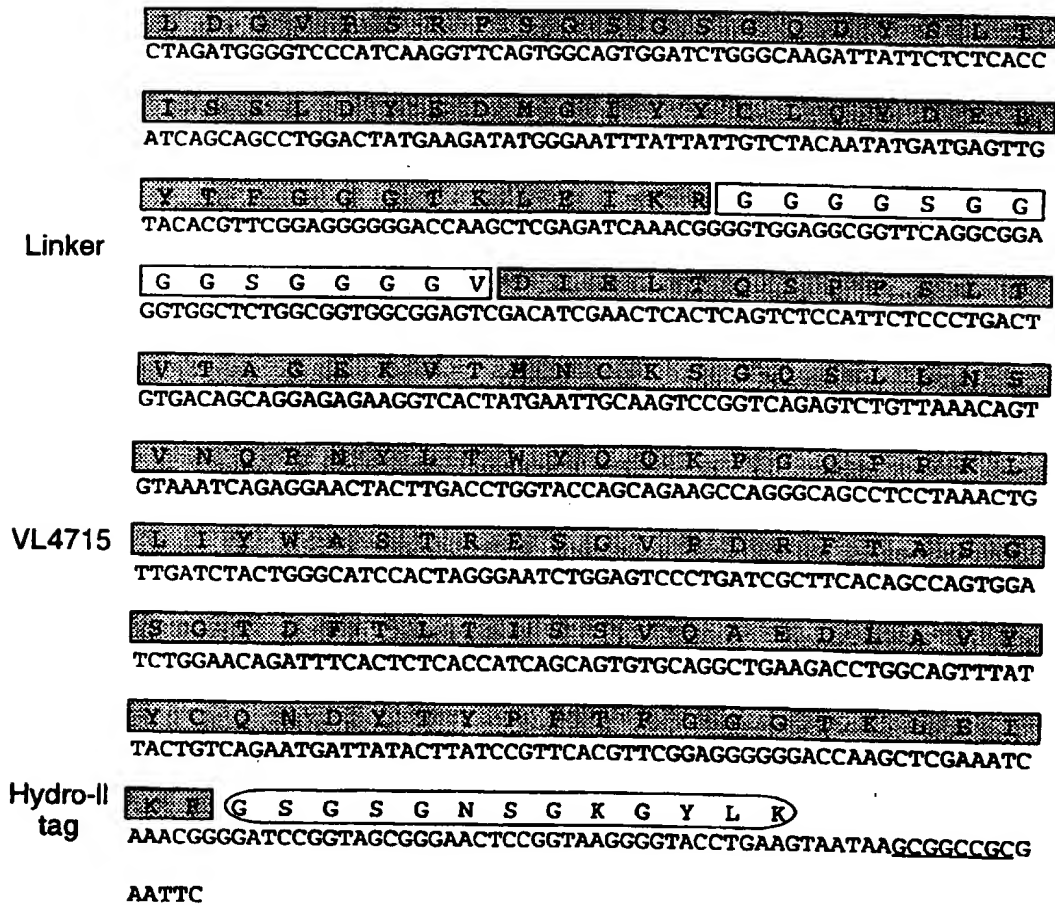
**VL3418** CTAGGAGAGAGAATCACTATCACTTGCAAGGCGAGTCAGGACATTAATACCTATTTAACC

W F Q Q K P G K S P R T L I Y R A N R T

TGGTTCCAGCAGAAACCAGGGAAATCTCCAAGACCCTGATCTATCGTGCAAACAGATTG

4/35

Fig.2B.



5/35

Fig.3A.

M K Y L L P T

AAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACG

pelB leader A A A G L L L L A A Q P A M A O V O L G  
 GCAGCCGCTGGATTGTTATTACTCGCTGCCCAACCAGCGATGGCCCAGGTGCAGCTGCAG

E S G P G L V A P S Q S L S I T C T V S  
 GAGTCAGGACCTGGCCTGGTGGCGCCCTCACAGAGCCTGTCCATCACATGCACCGTCTCA

G F S L T G L G V N H V R O P P G E G R  
 GGGTTCTCATTAAACCGCTATGGTGTAACCTGGGTTCCGACGCTCCAGGAAAGGGTCTG

E W L G M T W G D G N T D Y N S A L K S  
 GAGTGGCTGGGAATGATTGGGGTGATGGAAACACAGACTATAATTAGCTCTCAAATCC

VHlys R L S Y S K D N S A S Q V E L R M N S L  
 AGACTGAGCATCAGCAAGGACAACCTCCAAGAGCCAAGTTTTCTTAAAAATGAACAGTCTG

H T D D T A R Y Y C A H E R D Y R L D Y  
 CACACTGATGACACAGCCAGGTACTACTGTGCCAGAGAGAGAGATTATAGGCTTGACTAC

N G E G T T V T V S S G G G G S G G G G  
 TGGGGCGAAGGCACCAACGGTCAACGCTCTCCTCAGGTGGAGGCGGTTCAAGCGGAGGTGGC

Linker S G G G G S D I E L T Q S P A S L S A S  
 TCTGGCGGTGGCGGATCGGACATCGAGCTCACCCAGTCTCCAGCCTCCCTTTCTGCGTCT

V G E T V T T T C H A S G N T H N Y L A  
 GTGGGAGAACTGTCAACATCACATGTCGAGCAAGTGGGAATATTACAAATTATTAGCA

N Y D Q K R G E S E Q L L L V Y Y T T T D  
 TGGTATCAGCAGAAACAGGGAAATCTCCTCAGCTCCTGGTCTATTATACAACAACCTTA

A D G V P S R P S G S G S G T P Q Y S L P  
 GCAGATGGTGTGCCATCAAGGTTCAAGTGGCAGTGGATCAGGAACACAATATTCTCTCAAG

VLlys I N S L Q P E D E G S Y Y C Q H F W S T  
 ATCAACAGCCTGCAACCTGAAGATTTTGGGAGTTATTACTGTCAACATTTTGGAGTACT

P R T P L G G T K L E I K H E Q K L I S  
 CCTCGGACGTTGGTGGAGGCACCAAGCTCGAGATCAAACGGGAACAAAACTCATCTCA

Myc-tag E E D L N  
 GAAGAGGATCTGAATTAATAAGATCAAACGGTAATAAGGATCCAGCTCGAATTC

6/35

Fig.3B.

M K Y L L P T

AAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACG

pelB A A A G L L L L A A Q P A M A D V Q L D

leader GCAGCCGCTGGATTGTTATTACTCGCTGCCCAACCAGCGATGGCCCAGGTGCAGCTGCAG

E S G G D L V K P G G S L T E S C A M S

GAGTCAGGGGAGACTTAGTGAAGCCTGGAGGGTCCCTGACACTCTCCTGTGCAACCTCT

G F L R E E Y A F S W V D O T S D R S L

GGATTCACTTTCAGTAGTTATGCCTTTTCTTGGGTCCGCCAGACCTCAGACAAGAGTCTG

E W V A T T S S T D T Y T Y Y S E N V E

GAGTGGTTCGCAACCATCAGTAGTACTGATACTTATACCTATTATTTCAGACAATGTGAAG

VH4715 G R P T I S R E N G R N T L Y L Q M S S

GGCGCTTCACCATCTCCAGAGACAATGGCAAGAACACCCTGTACCTGCAAATGAGCAGT

L R S E D T A V Y N C A R H G Y Y G K G

CTGAAGTCTGAGGACACAGCCGTGTATTACTGTGCAAGACATGGGTACTATGGTAAAGGC

Y F D Y W G Q G T T V T V S S G G G G S

TATTTTGACTACTGGGGCCAAGGGACCACGGTCACCGTCTCCTCAGGTGGAGGCGGTTCA

Linker G G G G S G G G G S D I E L T Q S P P S

GGCGGAGGTGGCTCTGGCGGTGGCGGATCGGACATCGAGCTCACTCAGTCTCCATTCTCC

L T V T A G E K V T N N C K S G Q C L L

CTGACTGTGACAGCAGGAGAGAAGGTCACTATGAATTGCAAGTCCGGTCAGAGTCTGTTA

N S V N O R N Y L T R Y Q Q K P G Q P P

AACAGTGTAAATCAGAGGAACACTTGACCTGGTACCAGCAGAAGCCAGGGCAGCCTCCT

K L E I Y W A T S T R E S G Y P D R F T A

VL4715 AAAGTGTGATCTACTGGGCATCCACTAGGGAATCTGGAGTCCCTGATCGCTTCACAGCC

S G S G T D P T L T L S E V Q A E D D A

AGTGGATCTGGAACAGATTTCACTCTCACCATCAGCAGTGTGCAGGCTGAAGACCTGGCA

V Y Y C Q N D Y T Y P P T F G G E T K L

GTTTATTACTGTCAGAATGATTATACTTATCCGTTACGTTCCGAGGGGGGACCAAGCTC

Hydro2-tag E I K R G S G S G N S G K G Y L K

GAGATCAAACGGGGATCCGGTAGCGGGAACCTCCGGTAAGGGGTACCTGAAGTAATAAGAT

CAAACGGTAATAAGGATCCAGCTCGAATTC

7/35

Fig.4.

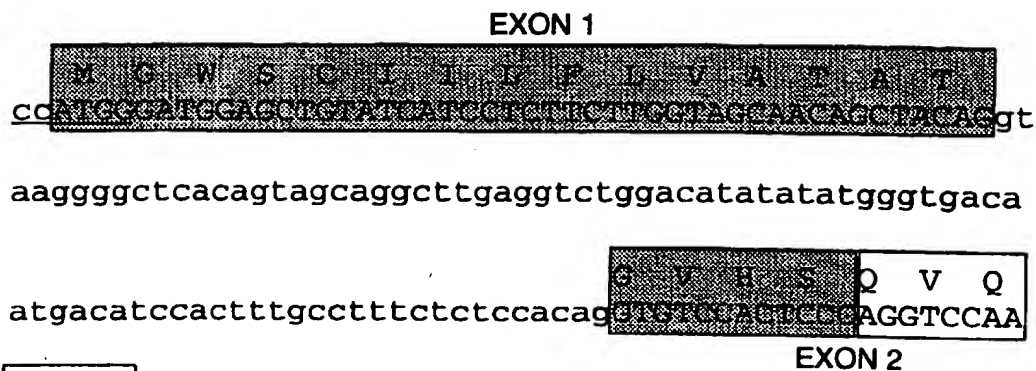
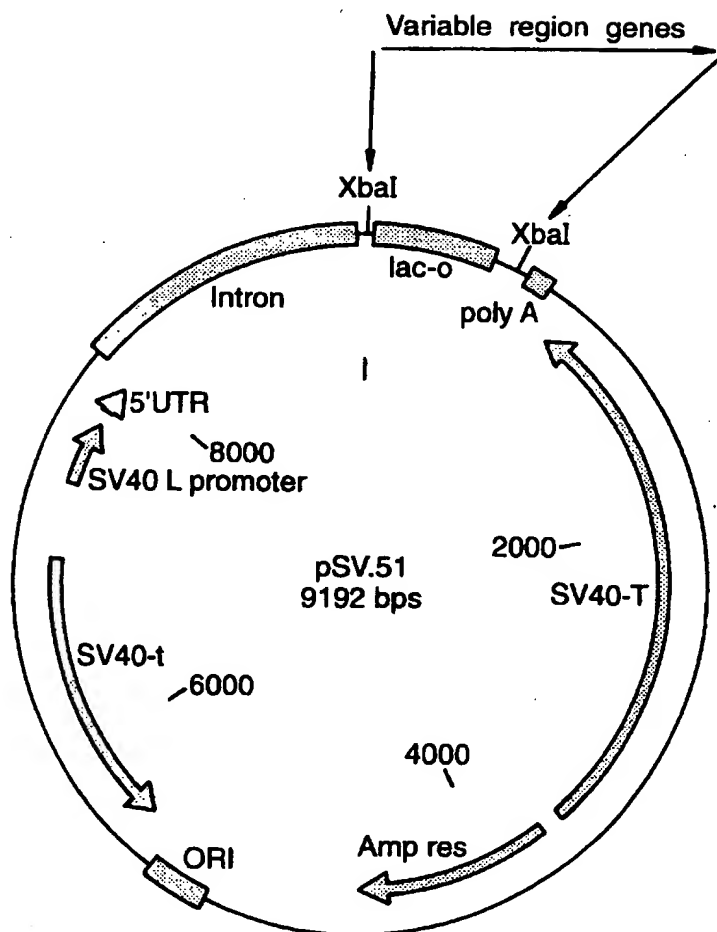


Fig.5.



8/35

Fig.6A.

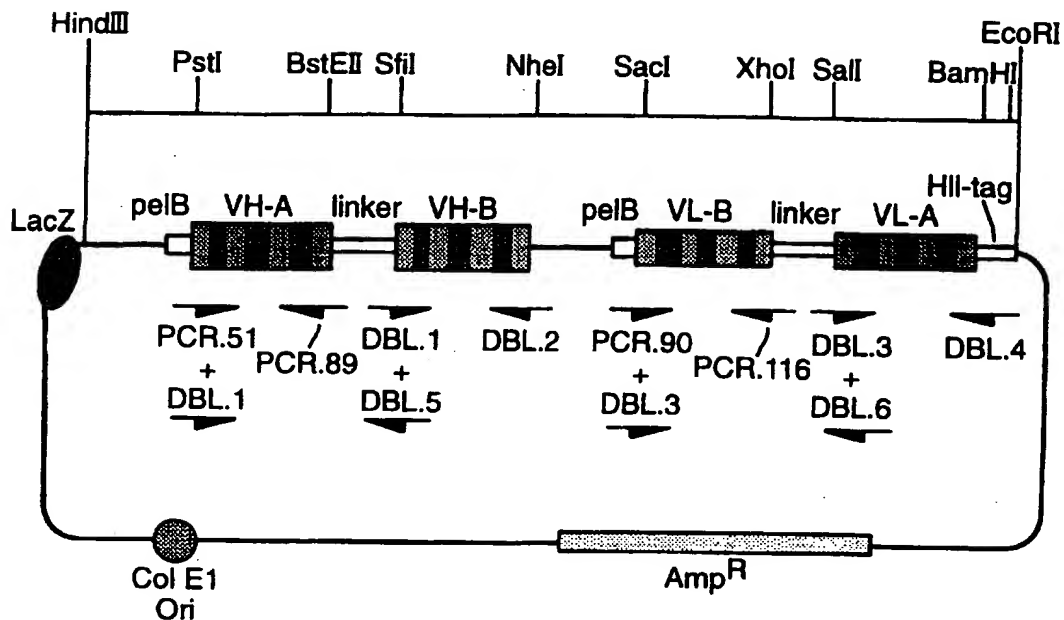
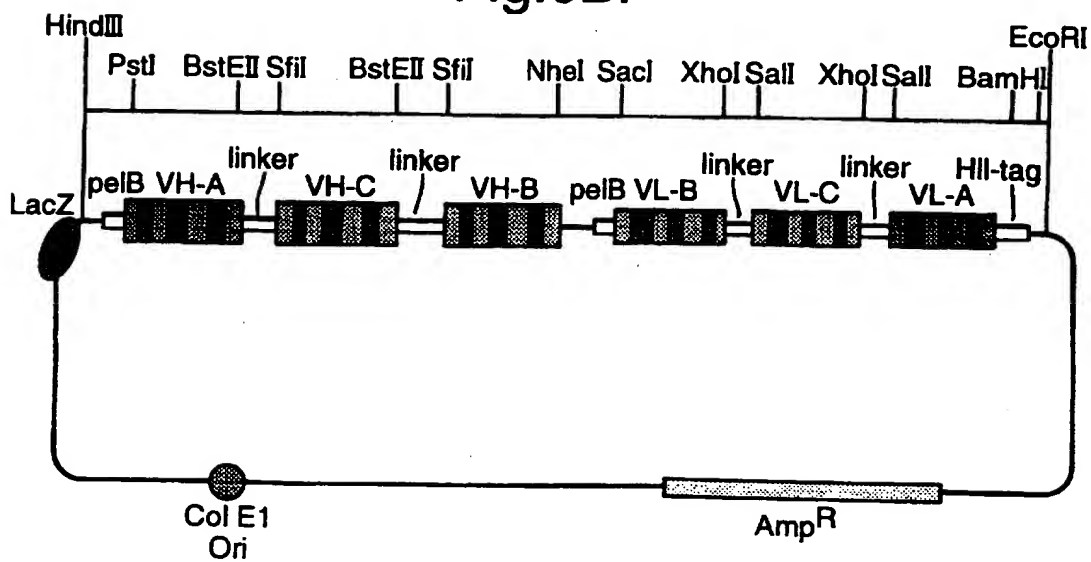
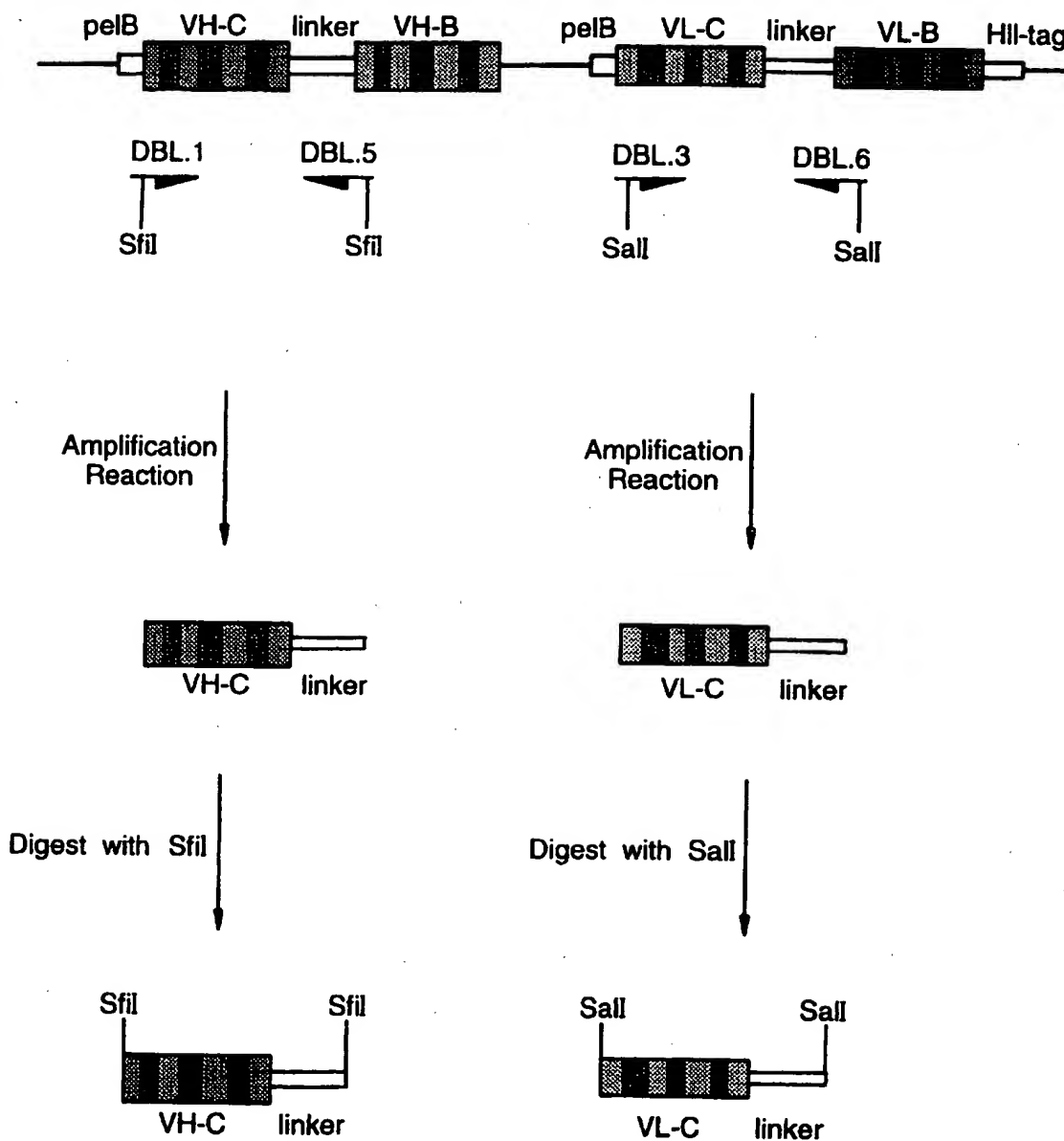


Fig.6B.



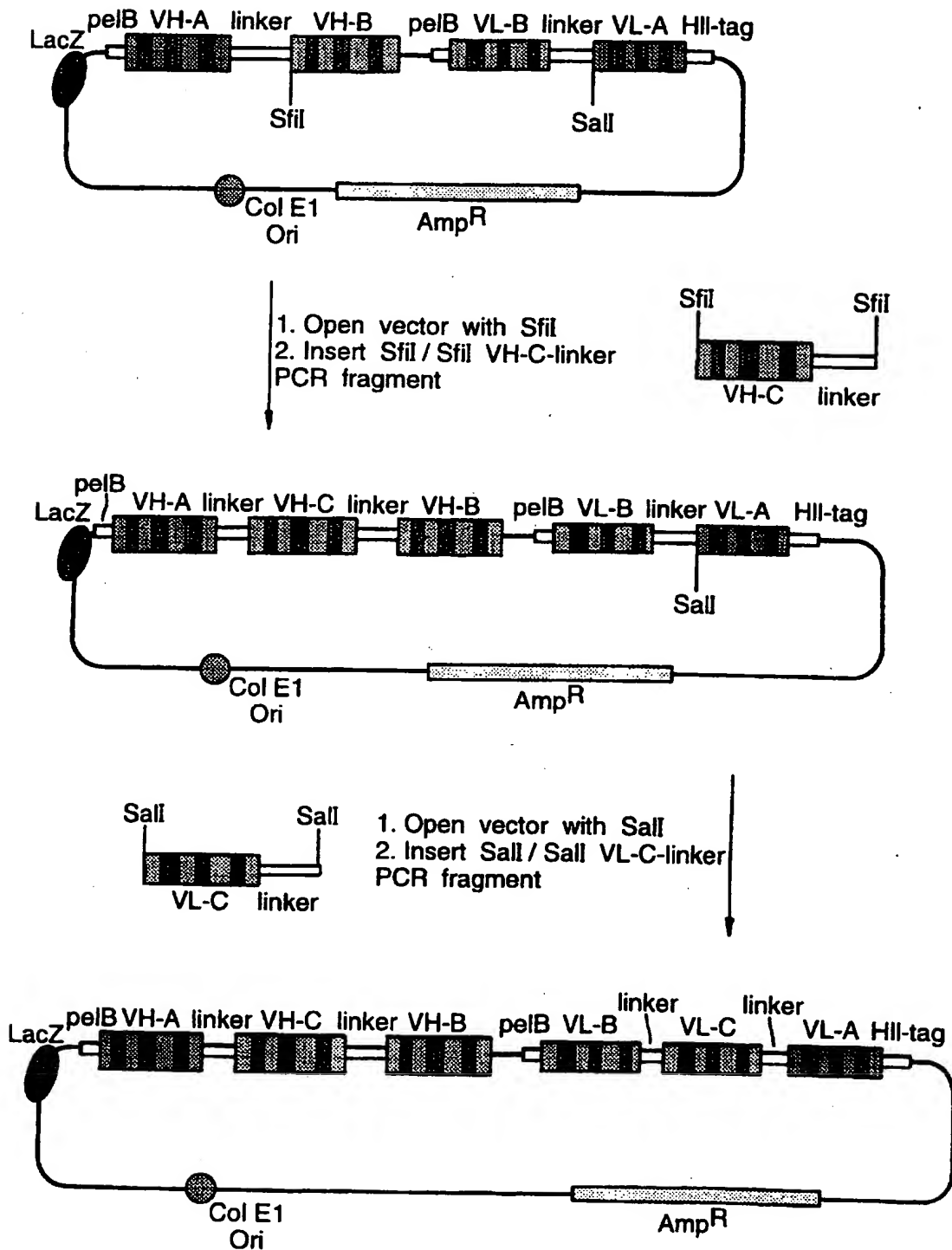
9/35

Fig.7A.



10/35

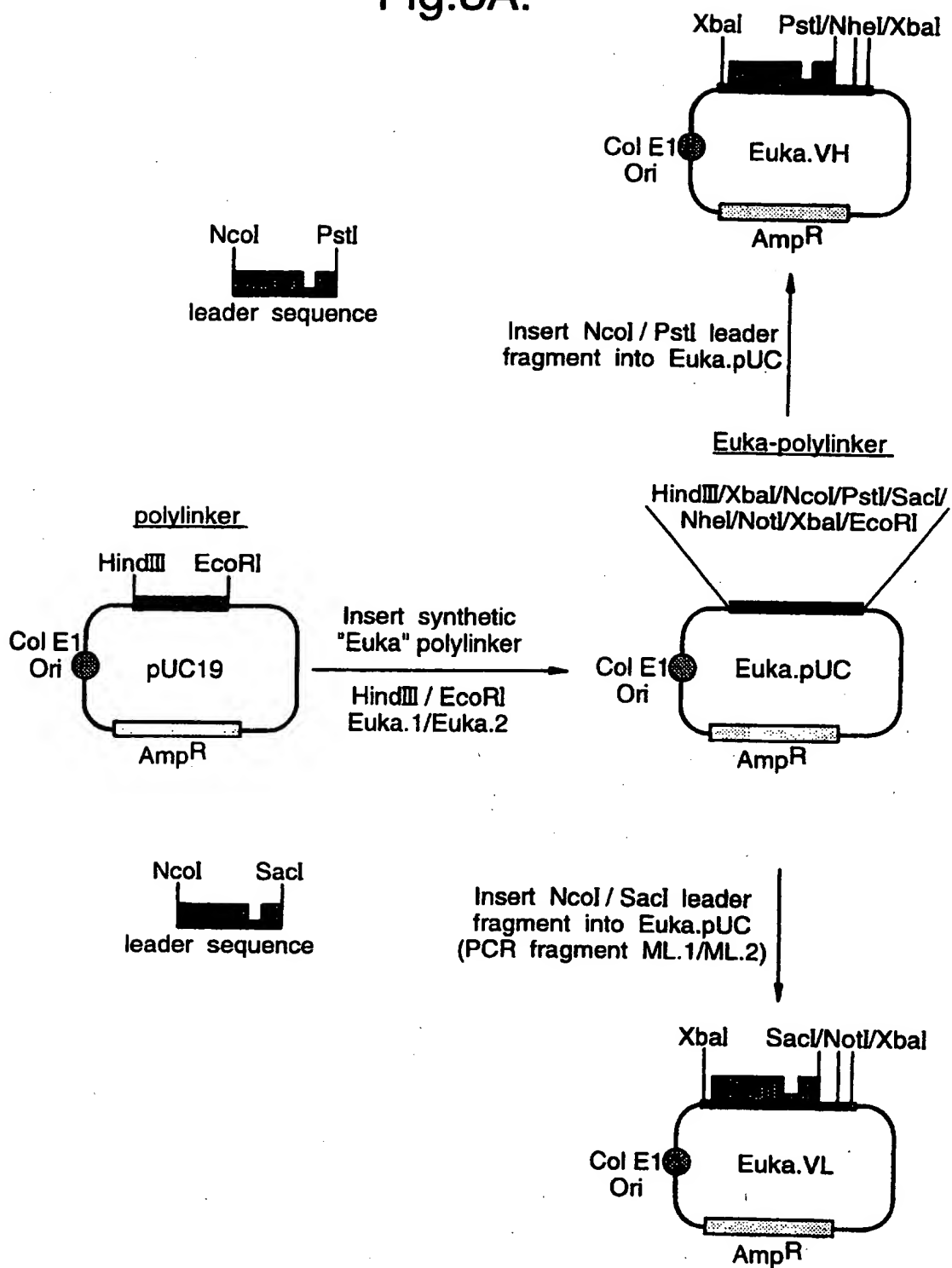
Fig.7B.





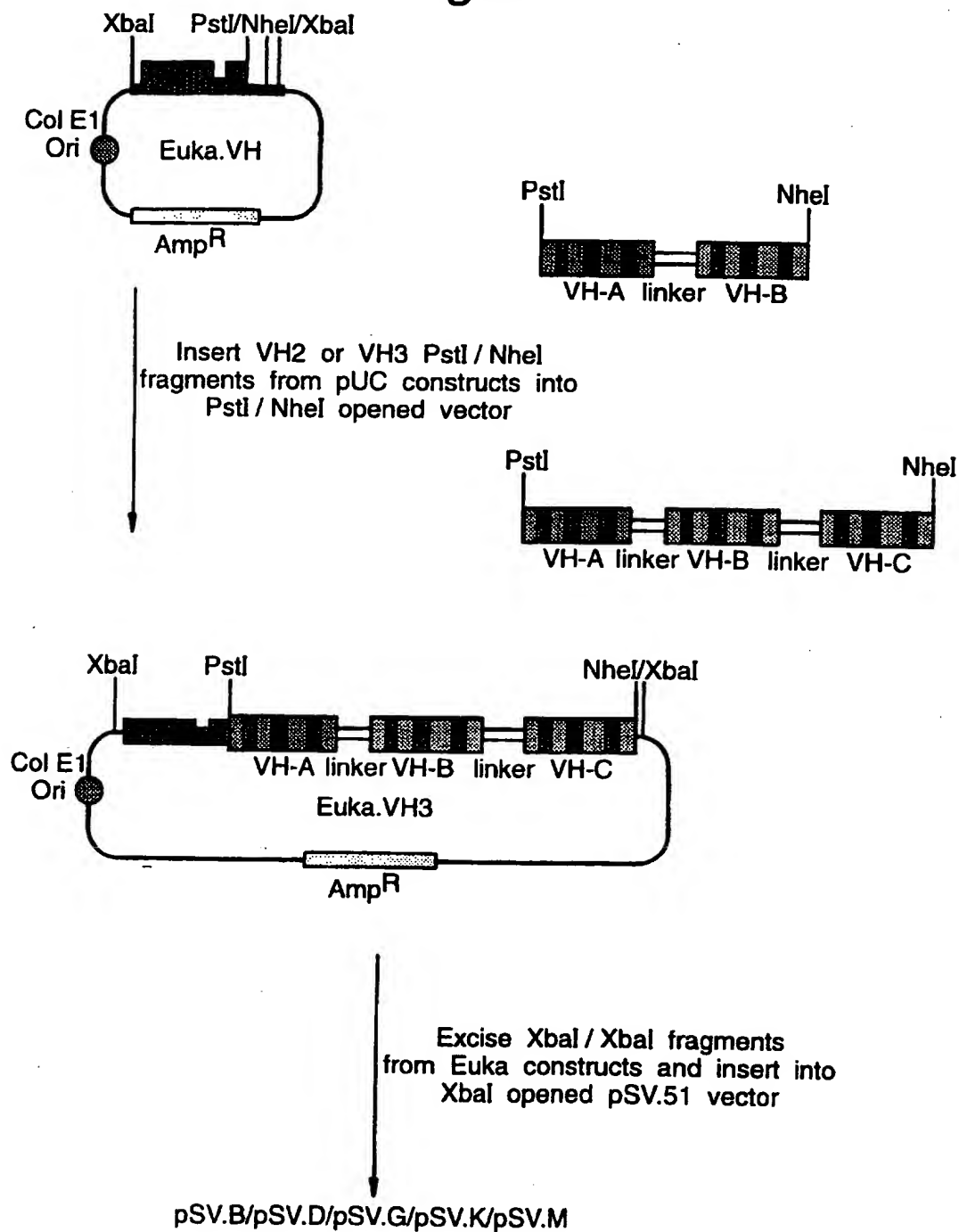
11/35

Fig.8A.



A35

Fig.8B.



13/35

Fig.8C.

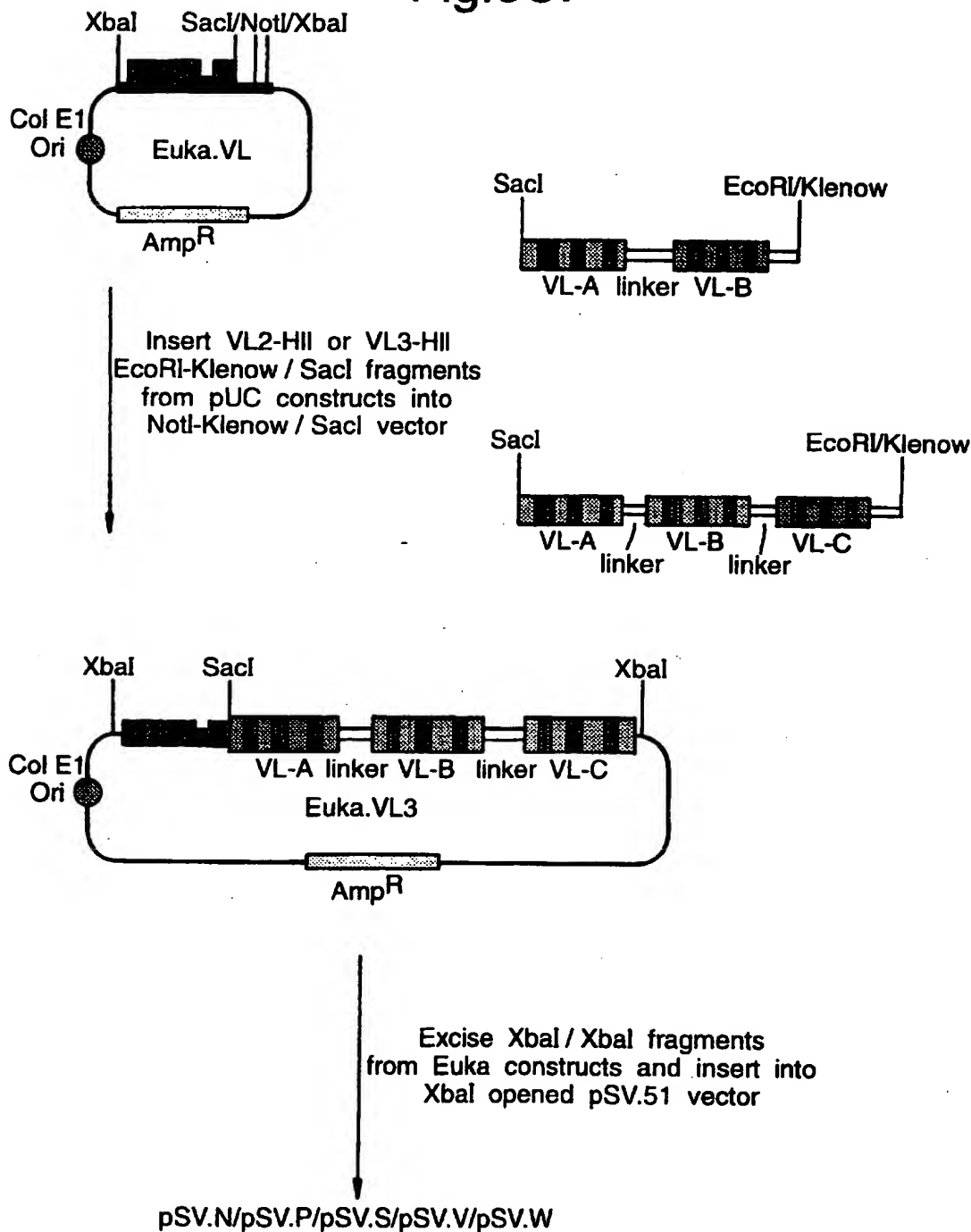
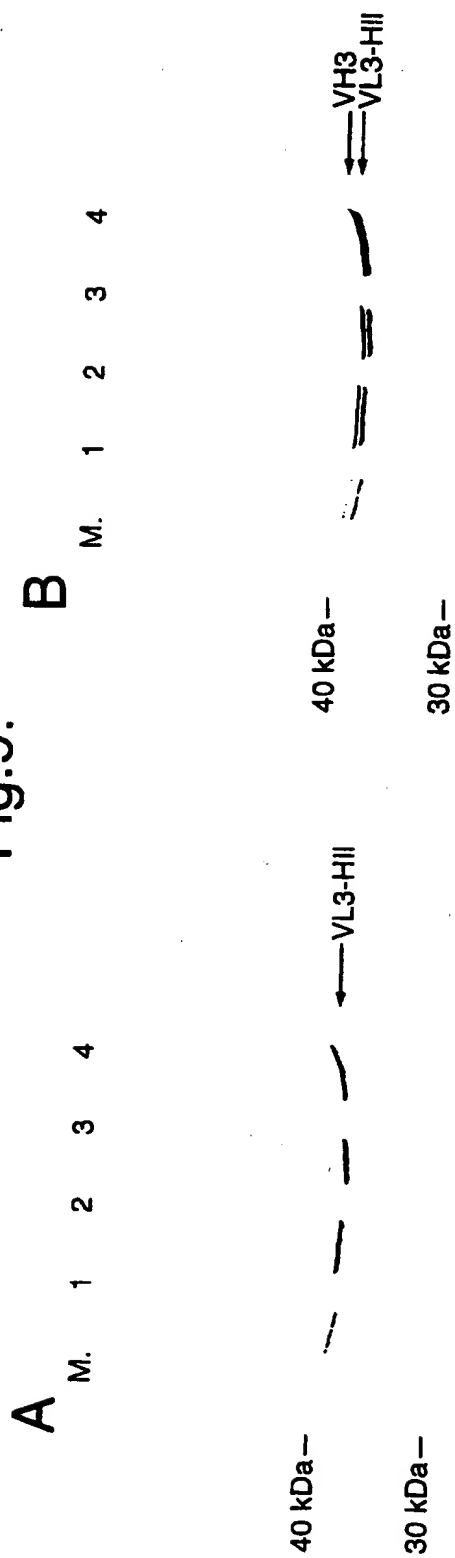
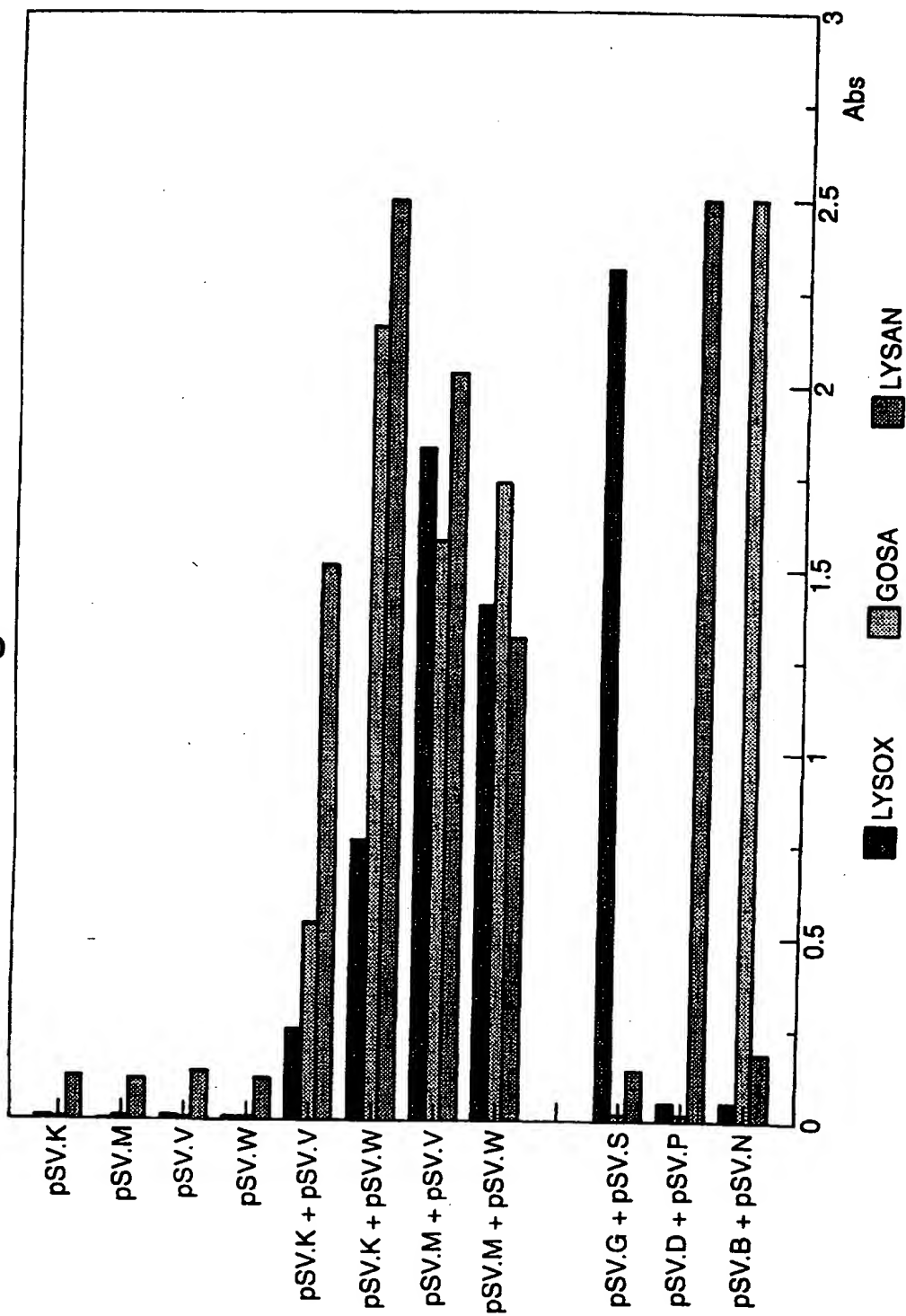


Fig.9.



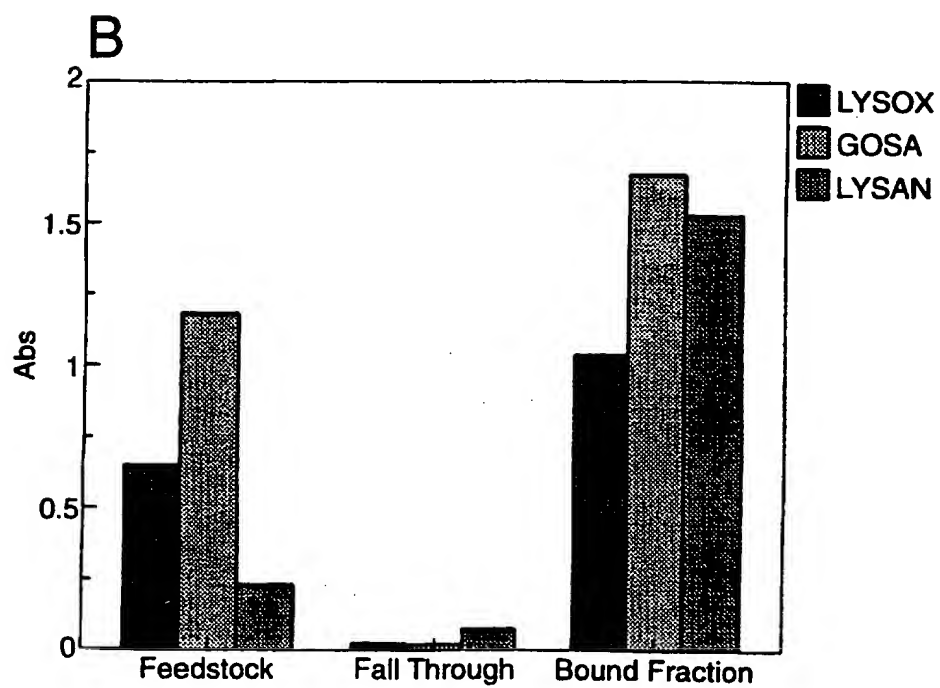
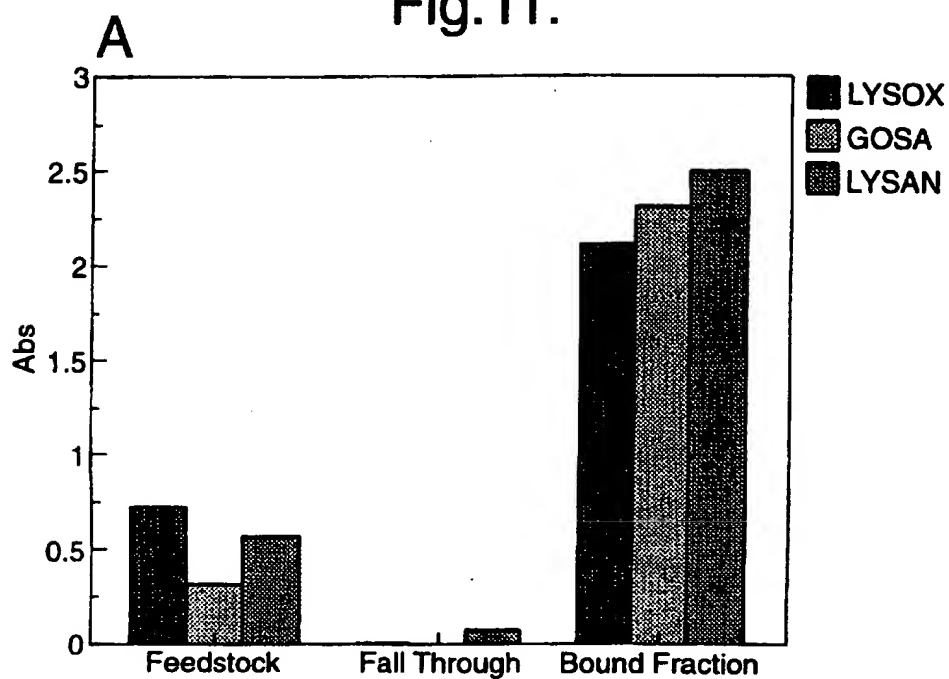
15/35

Fig.10.



10/35

Fig.11.



17/35

Fig.12.

D I E L T Q S F A S I S A S V G F  
 GAATTGGGCCGACATC GAGCTC ACCCAGTCTCCAGCCTCCCTTTCTGCGTCTGTGGGAGA  
 T V T I T C R A S G N I H N Y D A R Y C  
 AACTGTCACCATCACATGTCGAGCAAGTGGGAATATTCACAATTATTTAGCATGGTATCA  
 Q K Q G K S P Q I L V Y Y T T T L A D G  
 GCAGAAACAGGGAATCTCCTCAGCTCCTGGTCTATTATACAACAACCTTAGCAGATGG  
 V L Lys V P S R F S G S G S G T O Y S D R I N S  
 TGTGCCATCAAGGTTCAAGTGGCAGTGGATCAGGAACACAATATTCTCTCAAGATCAACAG  
 L Q P E D F G S Y Y C H F W S T P R T F  
 CCTGCAACCTGAAGATTTTGGGAGTTATTACTGTCAACATTTTGGAGTACTCCTCGGAC  
 G G G T K L E I K R G G G G S G G G G S  
 Linker GTTCGGTGGAAACCAAG CTCGAGATCA AACGGGTGGAGGCGGTTCAGGCGGAGGTGGCTC  
 G G G G S Q V D D Q E E G P G L V A P S  
 TGGCGGTGGCGGATCGCAGGTGCAG CTGCAGCAGT CAGGACCTGGCCTGGTGGCGCCCTC  
 Q S L S I T C T V S G E S L T G Y G V N  
 ACAGAGCCTGTCCATCACATGCACCGTCTCAGGGTTCTCATTAAACCGCTATGGTGTA  
 H V R O P P G E G L E W L G N F W G D G  
 CTGGGTTGCCAGCCTCCAGGAAGGGTCTGGAGTGGCTGGGAATGATTTGGGGTGATGG  
 N T D Y N S A L K E F L S I L E K D N E K  
 VHLys AAACACAGACTATAATTAGCTCTCAAATCCAGACTGAGCATCAGCAAGGACAACCTCAA  
 S I Q V F E E K M N S I H T D D T A R Y Y C  
 GAGCCAAGTTTCTTAAAAATGAACAGTCTGCACACTGATGACACAGCCAGGTACTACTG  
 A R E F D Y R L D Y W G Q G T H V T V E  
 TGCCAGAGAGAGAGATTATAGGCTTGACTACTGGGGCCAAGGGACACGGTCACCGTCTC  
 CTCATGATAAGCTT

18/35

Fig.13.

M K Y L L P T A  
 pelB AAGCTTGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAG  
 leader A A G L L L L A A Q P A M A Q V Q L Q Q  
 CCGCTGGATTGTTATTACTCGCTGCCCAACCAGCGATGGCCCAGGTGCAGCTGCAGCAGT  
 S G P E D V L P G A S V T M S C K A S G  
 CAGGACCTGAGCTGGTAAAGCCTGGGGCTTCAGTGAAGATGTCCTGCAAGGCTTCTGGAT  
 V T P T E Y V M H N V K Q K R G Q S H E  
 ACACATTCAGCTAGCTATGTTATGCAGCTGGGTGAAACAGAAGCCTGGGCAGGGCCTTGAGT  
 VH3418 H I G Y T Y P Y N D G T T Y N S E K F P G  
 GGATTGGATATATTATCCTTACAATGATGGTACTAAGTACAATGAGAAGTTCAAAGGCA  
 K A T L T S D K S S S T A Y N E L S S I  
 AGGCCACACTGACTTCAGACAAATCCTCCAGCACAGCCTACATGGAGCTCAGCAGCCTGA  
 T S R D S A V Y Y C S R R P D V W G Q Q  
 CCTCTGAGGACTCTGCGGTCTATTACTGTTCAAGACGCTTTGACTACTGGGGCCAAGGGA  
 T T V T V S S  
 CCACGGTCACCGTCTCCTCATAATAAGAGCTATGGGAGCTTGCATGCAAATTCTATTTCA  
 M K Y L L P T A A A G L L L L  
 pelB AGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCG  
 leader A A Q P A M A D I E L T Q S F S E N Y A  
 CTGCCCAACCAGCGATGGCCGACATCGAGCTCACCCAGTCTCCATCTTCCATGTATGCAT  
 E L G E R I H I T F C K A S Q D I N T V I  
 CTCTAGGAGAGAGAATCACTATCACTTGCAAGGCGAGTCAGGACATTAATACCTATTAA  
 T W F Q Q P P G K B P T L I Y R A N E  
 CCTGGTTCCAGCAGAAACCAGGGAAATCTCCAAGACCCTGATCTATCGTGCAAACAGAT  
 VL3418 L E D G V F R R S G S G Q D Y S I  
 TGCTAGATGGGGTCCCATCAAGGTTCACTGGCAGTGGATCTGGGCAAGATTATTCTCTCA  
 E I S S L D Y E D M S T Y Y T L G Y D E  
 CCATCAGCAGCCTGGACTATGAAGATATGGGAATTTATTATTGTCTACAATATGATGAGT  
 I Y T F G G G T E L E T K R  
 TGTACACGTTCCGAGGGGGGACCAAGCTCGAGATCAAACGGTAATAATGATCAAACGGT  
 ATAAGGATCCAGCTCGAATTC



19/35

Fig.14.

M K Y L L P T  
 pelB leader AAGCTTGCATGCAAATTCATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACG  
 A A A G L L L L A A Q P A M A D V Q L Q  
 GCAGCCGCTGGATTGTTATTACTCGCTGCCCCAACCAGCGATGGCCAGGTGCAGCTGCAG  
E G G G D L V K P G G S L T L S C D T S  
 GAGTCAGGGGAGACTTAGTGAAGCCTGGAGGGTCCCTGACACTCTCCTGTGCAACCTCT  
G P T P E S Y A R S M V R D T S D T S M  
 GGATTCACTTTTCAGTAGTTATGCCTTTTCTTGGGTCCGCCAGACCTCAGACAAGAGTCTG  
E M V A T I S S T D T Y E Y S D N V K  
 VH4715 GAGTGGGTGCGAACCATCAGTAGTACTGATACTTATACCTATTATTTCAGACAATGTGAAG  
G R F T I S R D N G E N T L Y D Q M S S  
 GGGCGCTTCACCATCTCCAGAGACAATGGCAAGAACACCCTGTACCTGCAAATGAGCAGT  
L K S E D T A V Y Y C A R H E Y Y G K T G  
 CTGAAGTCTGAGGACACAGCCGTGTATTACTGTGCAAGACATGGGTACTATGGTAAAGGC  
Y F D Y W G Q G T T V T V S S  
 TATTTTGA TACTGGGGCCAAGGGACCACGGTCACTCTCCTCATAATAAGAGCTATGG  
 M K Y L L P T  
 pelB leader GAGCTTGCATGCAAATTCATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACG  
 A A A G L L L L A A Q P A M A D L E L T  
 GCAGCCGCTGGATTGTTATTACTCGCTGCCCCAACCAGCGATGGCCGACATCGAGCTCACT  
Q S P F E L T V T R C E R V T M N C R S  
 CAGTCTCCATTCTCCCTGACTGTGACAGCAGGAGAGAAGGTCACTATGAATTGCAAGTCC  
G D S L L N S V N Q R N Y L T D Y Q Q R  
 GGTCAGAGTCTGTAAACAGTGTAAATCAGAGGAACTACTTGACCTGGTACCAGCAGAAG  
P C Q P P K L L I Y W A S I R E S G V P  
 CCAGGGCAGCCTCCTAAACTGTTGATCTACTGGGCATCCACTAGGGAATCTGGAGTCCCT  
 VL4715 D H F T S S G G T D E F L T I S S V Q  
 GATCGCTTACAGCCAGTGGATCTGGAACAGATTCACTCTCACCATCAGCAGTGTGCAG  
A E D L A V Y Y C Q N D Y T Y E F T P C  
 GCTGAAGACCTGGCAGTTTATTACTGTGAGAATGATTATACTTATCCGTTACGTTCCGA  
 Myc-tag G G T K L E I K R E Q K L I S E E D L N  
 GGGGGGACCAAGCTCGAGATCAAACGGGAACAAAACTCATCTCAGAAGAGGATCTGAAT  
 TAATAAGATCAAACGGTAATAAGGATCCAGCTCGAATTC

20/35

Fig.15.

pelB leader AAGCTTGCATGCAAATTCATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACG  
 M K Y L L P T  
 A A A G L L L L A A Q P A M A Q V P L Q  
 GCAGCCGCTGGATTGTTATTACTCGCTGCCCAACCAGCGATGGCCAGGTGCAGCTGCAG  
 E S G G D L V K P G G S L T L S C A T S  
 GAGTCAGGGGAGACTTAGTGAAGCCTGGAGGGTCCCTGACACTCTCCTGTGCAACCTCT  
 G F T P S S V A F S W V E O T S D K S L  
 GGATTCACTTTCACTAGTTATGCCTTTTCTTGGGTCCGCCAGACCTCAGACAAGAGTCTG  
 E W V A T I S S T D T Y T Y Y S D N V K  
 VH4715 GAGTGGGTCGCAACCATCAGTAGTACTGATACTTATACCTATTATTTCAGACAATGTGAAG  
 G R F T I S R D N G K N T L Y L Q M S S  
 GGGCGCTTCACCATCTCCAGAGACAATGGCAAGAACACCCTGTACCTGCAAAATGAGCAGT  
 L K S E D T A V Y Y C A R H G Y Y G K G  
 CTGAAGTCTGAGGACACAGCCGTGTATTACTGTGCAAGACATGGGTACTATGGTAAAGGC  
 Y F D Y W G O G T T Y T V S S G G G G S  
 Linker TATTTTGACTACTGGGGCCAAGGGACCAGGTCAACCGTCTCCTCAGGTGGAGGCGGTTCA  
 G G G G S G G G G S D I E L T D E P F S  
 GGGGAGGTGGCTCTGGCGGTGGCGGATCGGACATCGAGCTCACTCAGTCTCCATCTCC  
 L T V T A D E K V T M N D K S G D S L E  
 CTGACTGTGACAGCAGGAGAGAAGGTCACTATGAATTGCAAGTCCGGTCAGAGTCTGTTA  
 N E V N O F R Y L T A Y Q G K P G D P P  
 VL4715 AACAGTGTAATCAGAGGAACACTTGACCTGGTACCAGCAGAAGCCAGGGCAGCCTCCT  
 K L L I Y W A S T R E S G V P D R Y T A  
 AAAGTGTGATCTACTGGGCATCCACTAGGGAATCTGGAGTCCCTGATCGCTTCACAGCC  
 G G S E T D F T D T L E S V D A E D L A  
 AGTGGATCTGGAACAGATTTCACTCTCACCATCAGCAGTGTGCAGGCTGAAGACCTGGCA  
 V Y Y C E D N D Y E T Y P P T F G O C T F E  
 GTTTATTACTGTGAGAATGATTATACTTATCCGTTACGTTCCGAGGGGGGACCAAGCTC  
 Myc-tag E K K R E Q K L I S E E D L N  
 GAGATCAAACGGGAACAAAACTCATCTCAGAAGAGGATCTGAATTAATAAGATCAAACG  
 GTAATAAGGATCCAGCTCGAATTC

21/35

## Fig.16A.

M K Y L L P T

AAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACG

pelB leader A A A G L L L L A A Q P A M A Q V Q L Q Q  
GCAGCCGCTGGATTGTTATTACTCGCTGCCCAACCAGCGATGGCCCAGGTGCAGCTGCAG

E S G G D L V K P G S S L F D S L A T S  
GAGTCAGGGGGAGACTTAGTGAAGCCTGGAGGGTCCCTGACACTCTCCTGTGCAACCTCT

G F T R S E Y A F S W L R O T S D K S L  
GGATTCACCTTTCAGTAGTTATGCCTTTTCTTGGGTCCGCCAGACCTCAGACAAGAGTCTG

E W V A T I S S L D T Y T Y Y S D R N V T  
GAGTGGGTGCGCAACCATCAGTAGTACTGATACTTATACCTATTATTTCAGACAATGTGAAG

VH4715 C R F T L S R D N G E N T L Y A Q N S S  
GGGCGCTTCACCATCTCCAGAGACAATGGCAAGAACACCCTGTACCTGCAAATGAGCAGT

L K S E D T A V L Y Y C A R H G Y Y G K G  
CTGAAGTCTGAGGACACAGCCGTGTATTACTGTGCAAGACATGGGTACTATGGTAAAGGC

Y F D Y W G Q G T T V T V S S G G G G S  
TATTTTGACTACTGGGGCCAAGGGACCACGGTCACCGTCTCCTCAGGTGGAGGCGGTTCA

Linker G G G G S G G G G S A G S A Q V Q L Q Q  
GGCGGAGGTGGCTCTGGCGGTGGCGGATCGGCCGGTTCGGGCCAGGTCCAGCTGCAACAG

S G P E L V I P G A S V T U S C K A S G  
TCAGGACCTGAGCTGGTAAAGCCTGGGGCTTCAGTGAAGATGTCTGCAAGGCTTCTGGA

Y T F T S Y V M H W V K Q K P G Q G L E  
TACACATTCAGTAGCTATGTTATGCACTGGGTGAAACAGAAGCCTGGGCAGGGCCTTGAG

VH3418 W I G Y T Y R Y N D G T K Y N E R P K G  
TGGATTGCATATATTTATCCTTACAATGATGGTACTAAGTACAATGAGAAGTTCAAAGGC

K A T L T S D K S S S T A Y M E L S S L  
AAGGCCACACTGACTTCAGACAAATCCTCCAGCACAGCCTACATGGAGCTCAGCAGCCTG

T S E D S A V Y Y C S R R F D Y W G Q G  
ACCTCTGAGGACTCTGCGGTCTATTACTGTTCAAGACGCTTTGACTACTGGGGCCAAGGG

T T V T V S S  
ACCACCGTCACCGTCTCCTCATAATAAGCTAGCGGAGCTGCATGCAAATTCTATTTCAAG

pelB leader M K Y L L P T A A A G L L L L A  
GAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCT

A Q P A M A D T E L T Q S P S S M T A S  
GCCCAACCAGCGATGGCCGACATCGAGCTCACCCAGTCTCCATCTTCCATGTATGCATCT

VL3418 L G E R L T I T C K A S Q D I N T Y L E  
CTAGGAGAGAGAATCACTATCACTTGCAAGGCGAGTCAGGACATTAATACCTATTTAACC

W F Q Q T P G K S P R L L Y R A N P E  
TGTTCCAGCAGAAACCAGGGAAATCTCCAAGACCCTGATCTATCGTGCAACAGATTG

Fig.16B.

Linker

VL4715

L D G V F S E F S G S G S G D Y S I T  
CTAGATGGGGTCCCATCAAGGTTCAAGTGGCAGTGGATCTGGGCAAGATTATTCTCTCACC

I E S L D V E D M G L Y Y C D Q Y E E F  
ATCAGCAGCCTGGACTATGAAGATATGGGAATTTATTATTGTCTACAATATGATGAGTTG

Y T F G G G T K L E I K R G G G G S G G  
TACACGTTCCGAGGGGGGACCAAGCTCGAGATCAAACGGGGTGGAGGCGGTTCAAGCGGA

G G S G G G G V D I E L T Q S P F G E T  
GGTGGCTCTGGCGGTGGCGGAGTCGACATCGAACTCACTCAGTCTCCATTCTCCCTGACT

V T A G E E V Y T I N C K S G Q E L N E  
GTGACAGCAGGAGAGAAGGTCACTATGAATTGCAAGTCCGGTCAGAGTCTGTTAAACAGT

V N O R N Y L T W Y D Q K P G G P P K L  
GTAAATCAGAGGAACACTTGACCTGGTACCAGCAGAAGCCAGGGCAGCCTCCTAAACTG

L I Y H A S T R E S G V P D E F T A S G  
TTGATCTACTGGGCATCCACTAGGGAATCTGGAGTCCCTGATCGCTTCACAGCCAGTGA

S G D D E T E T I E S V Q A E D I A V F  
TCTGGAACAGATTTCACTCTCACCATCAGCAGTGTGCAGGCTGAAGACCTGGCAGTTTAT

Y C D N D V Y P F T P G G G T K L E I  
TACTGTCAGAATGATTATACTTATCCGTTACGTTCCGAGGGGGGACCAAGCTCGAAATC

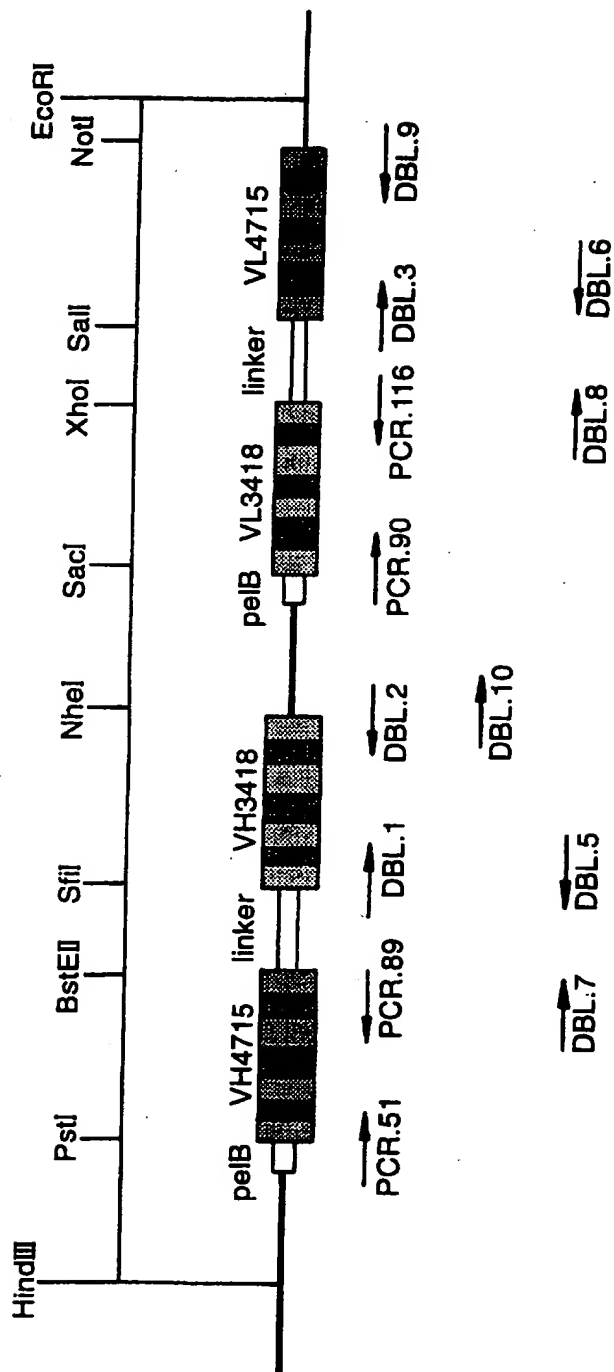
K R

AAACGGTAATAAGCGCCGGAATTC

23/35

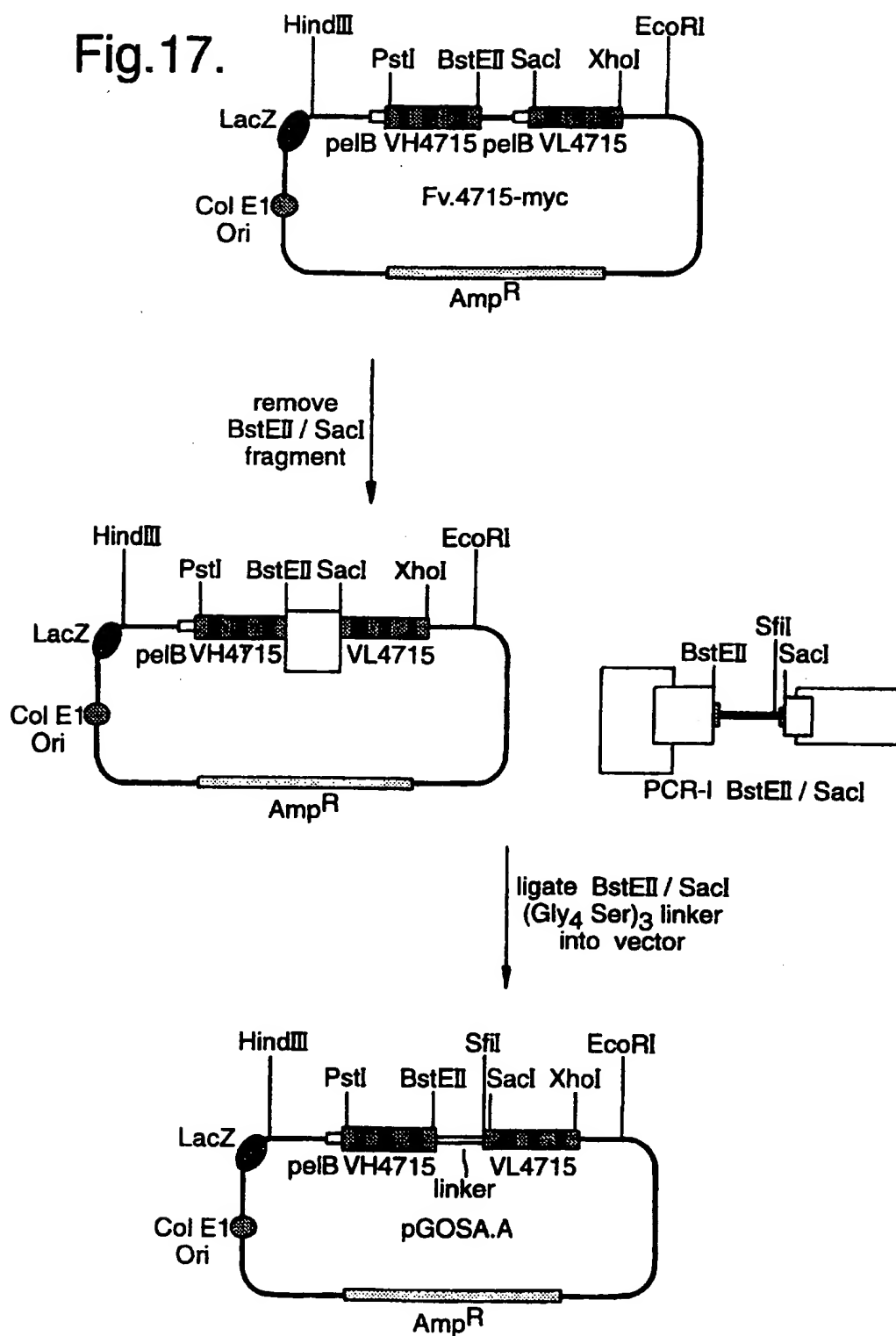
Fig. 16C.

pgOSA.E



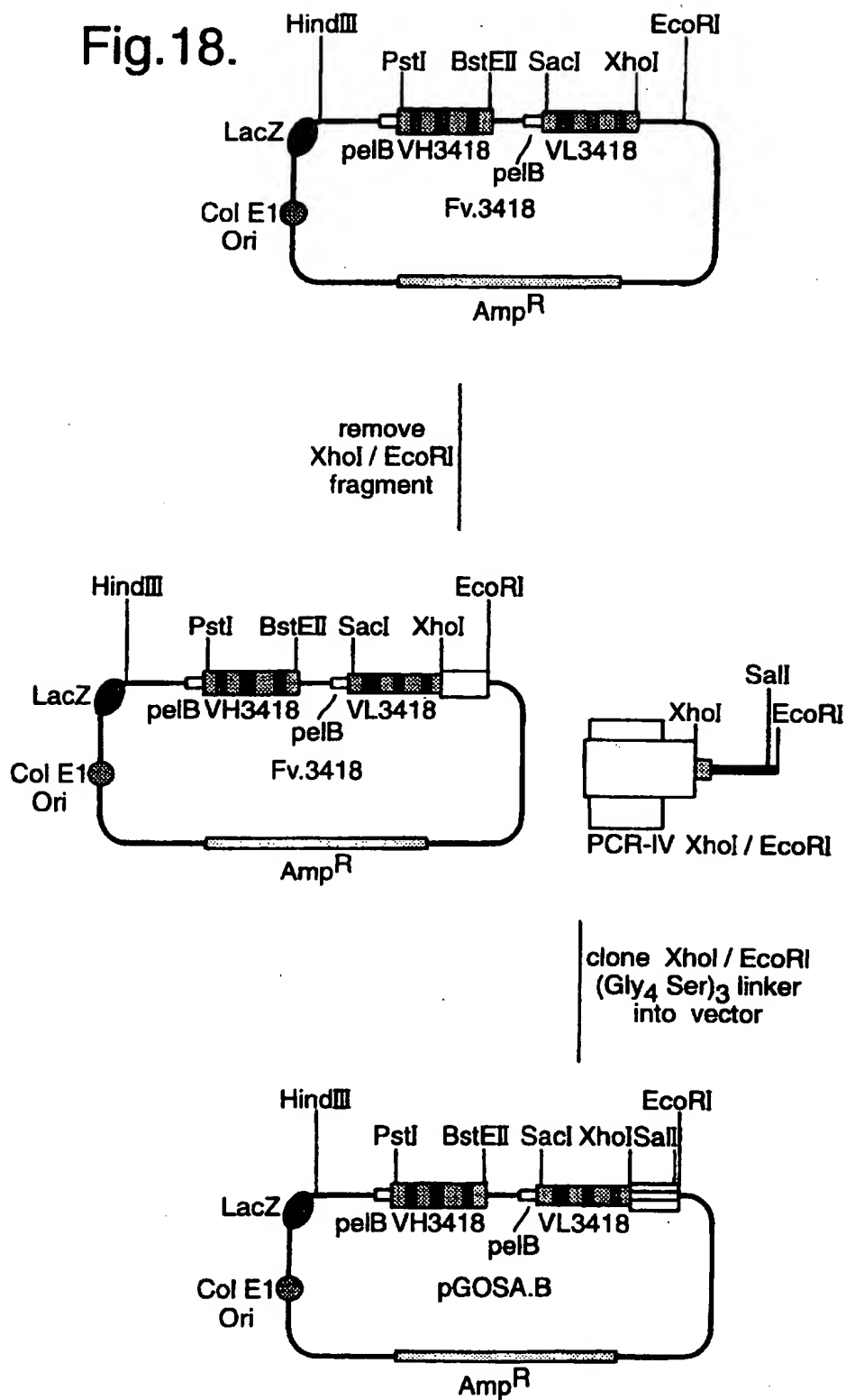
24/35

Fig.17.



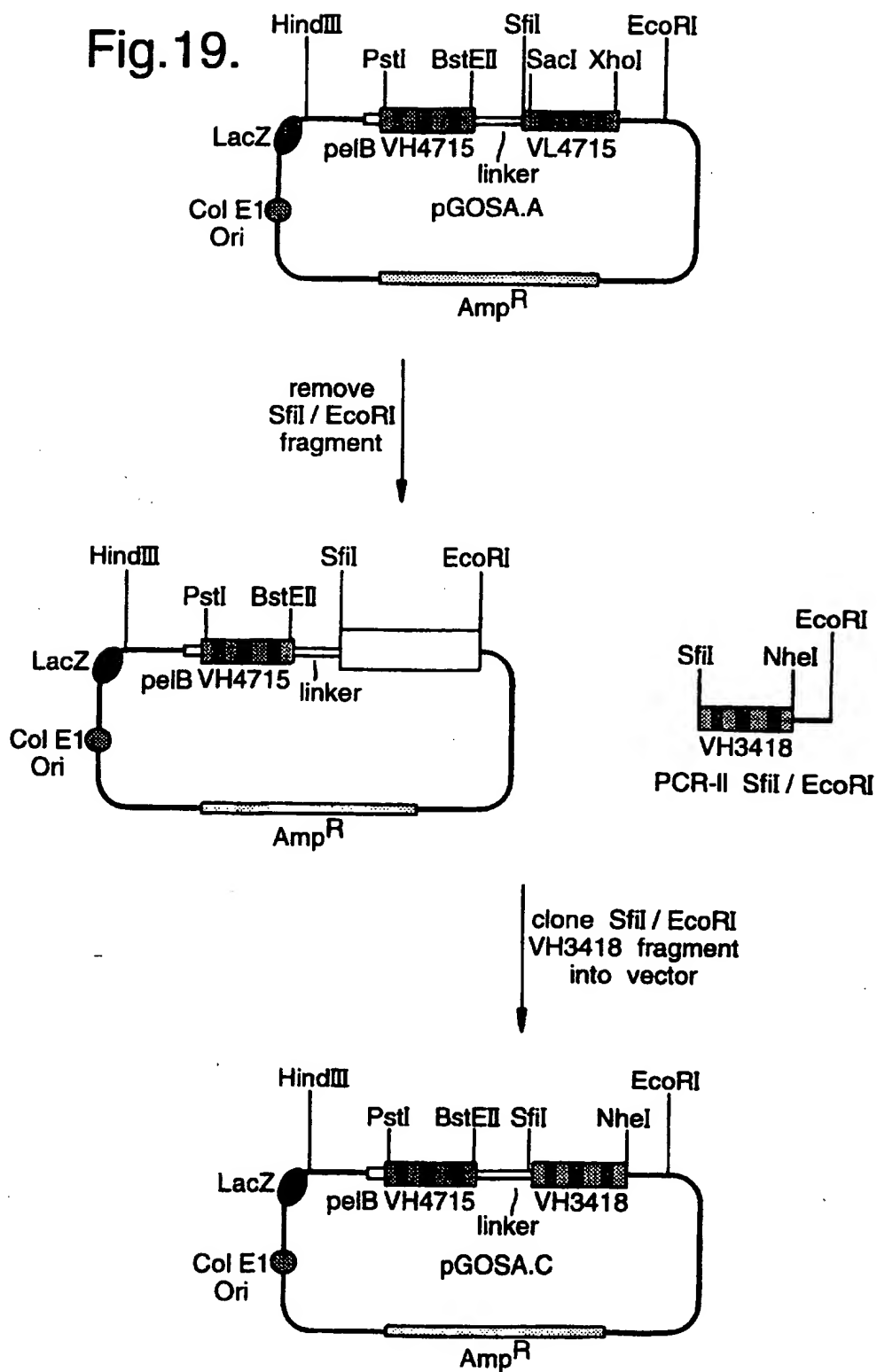
25/35

Fig.18.



26/35

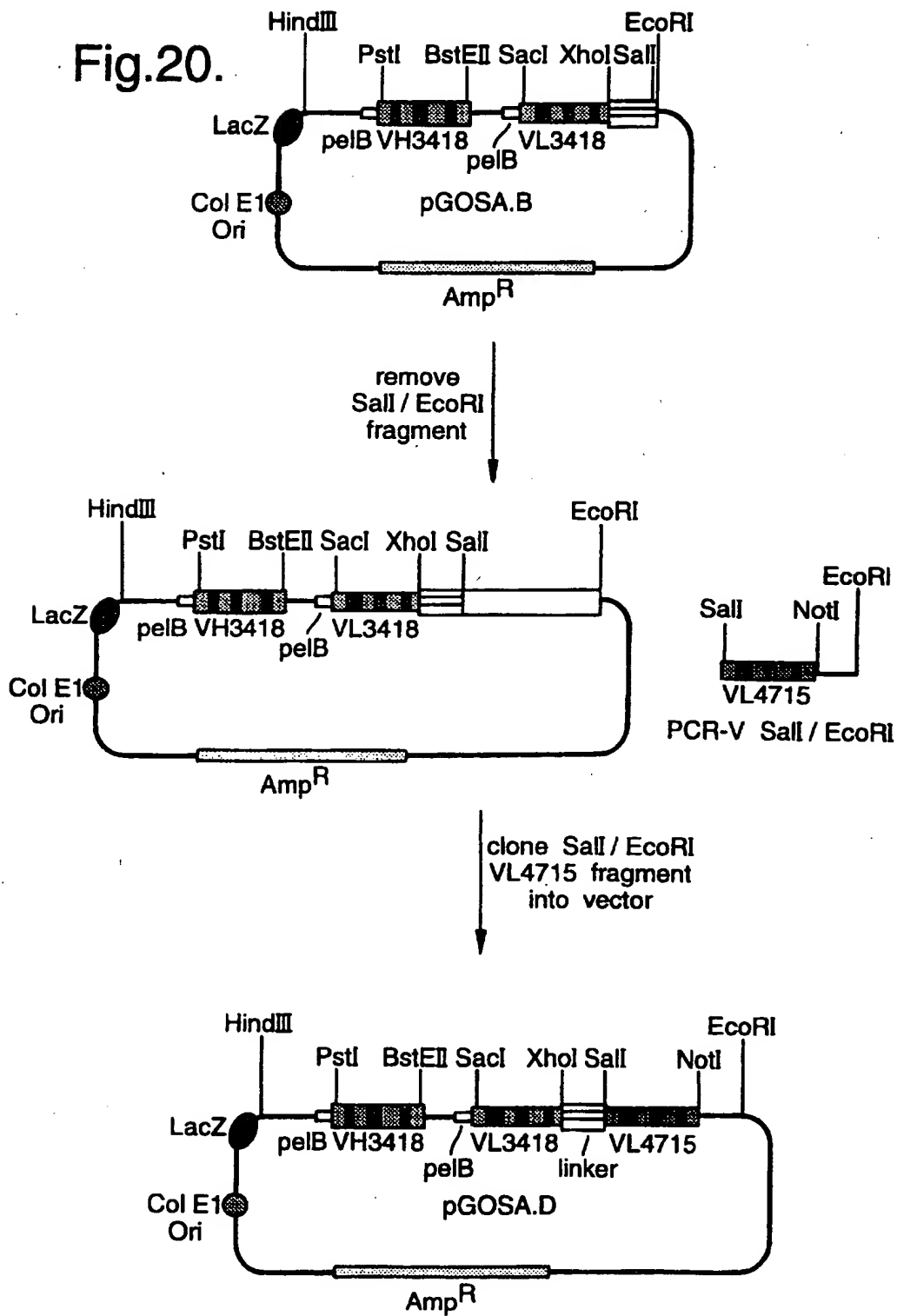
Fig.19.





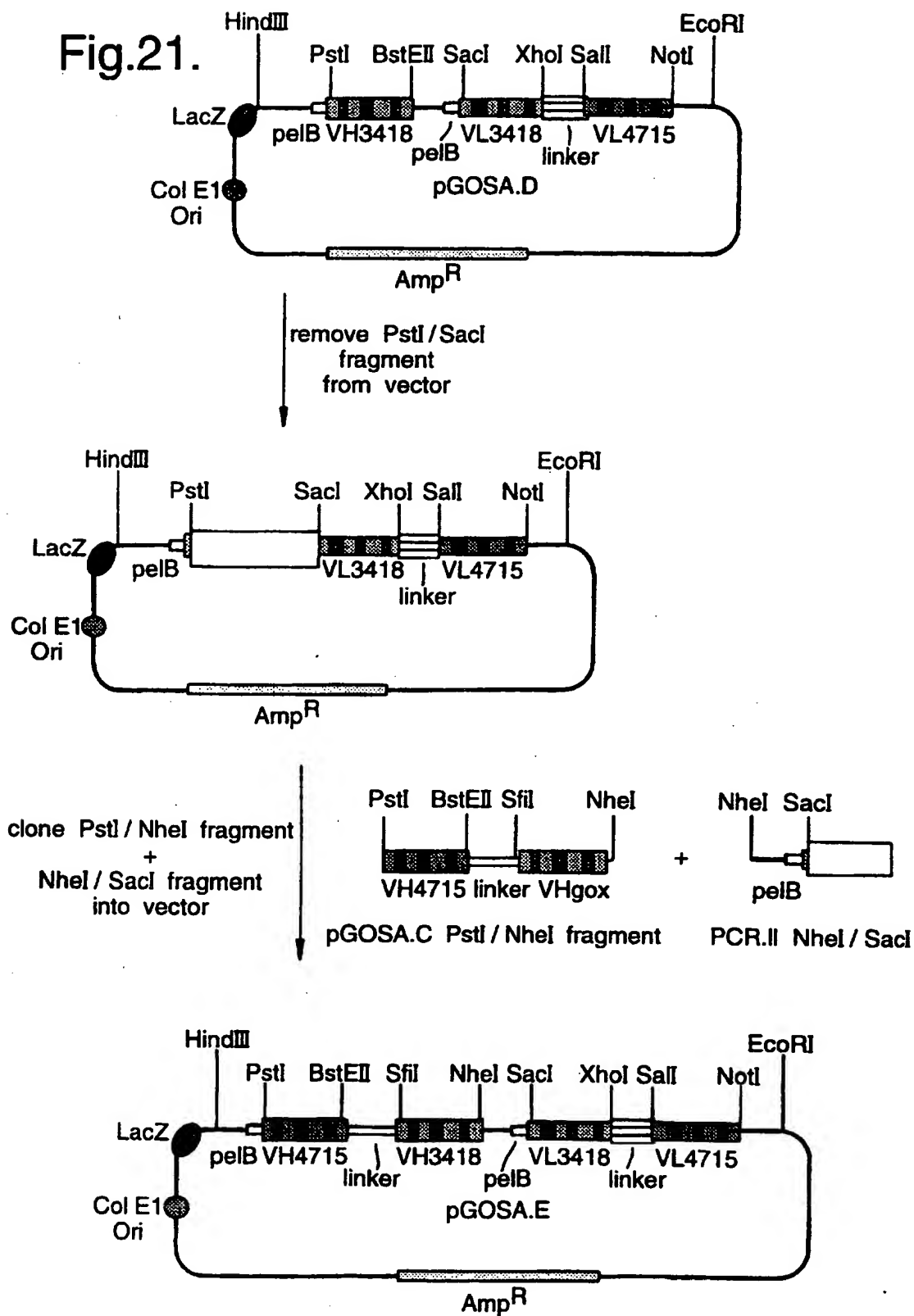
27/35

Fig.20.



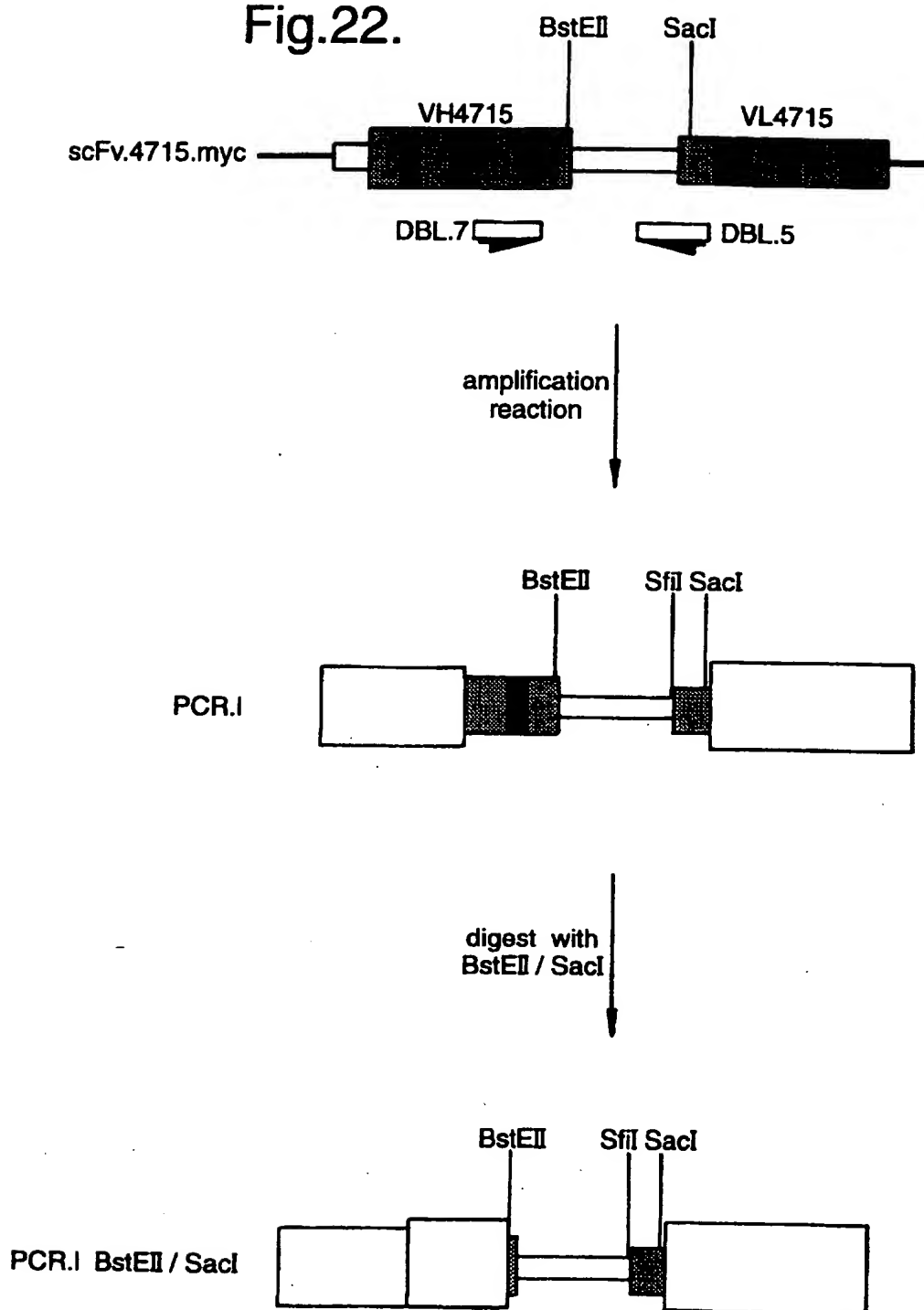
28/35

Fig.21.



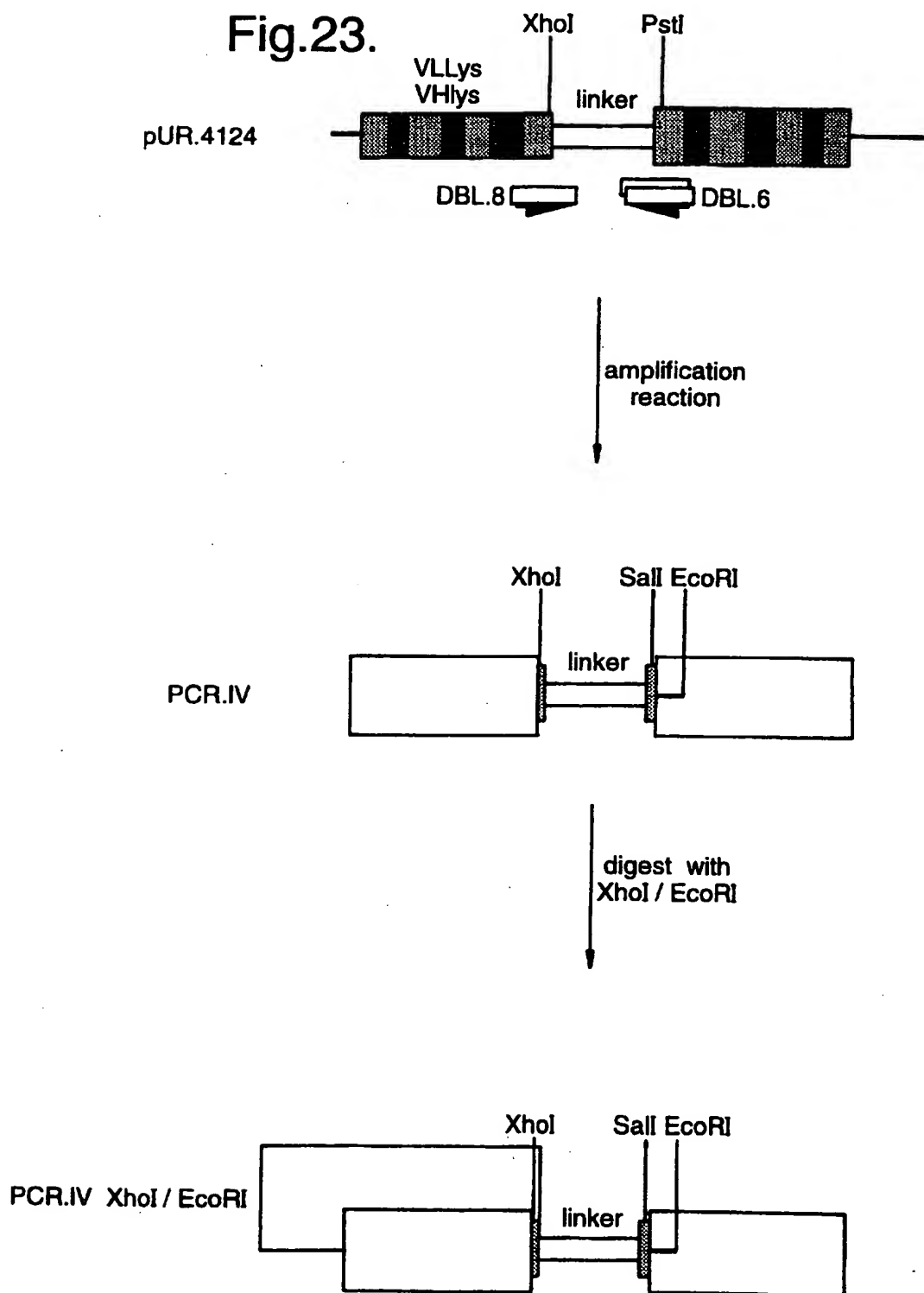
29/35

Fig.22.



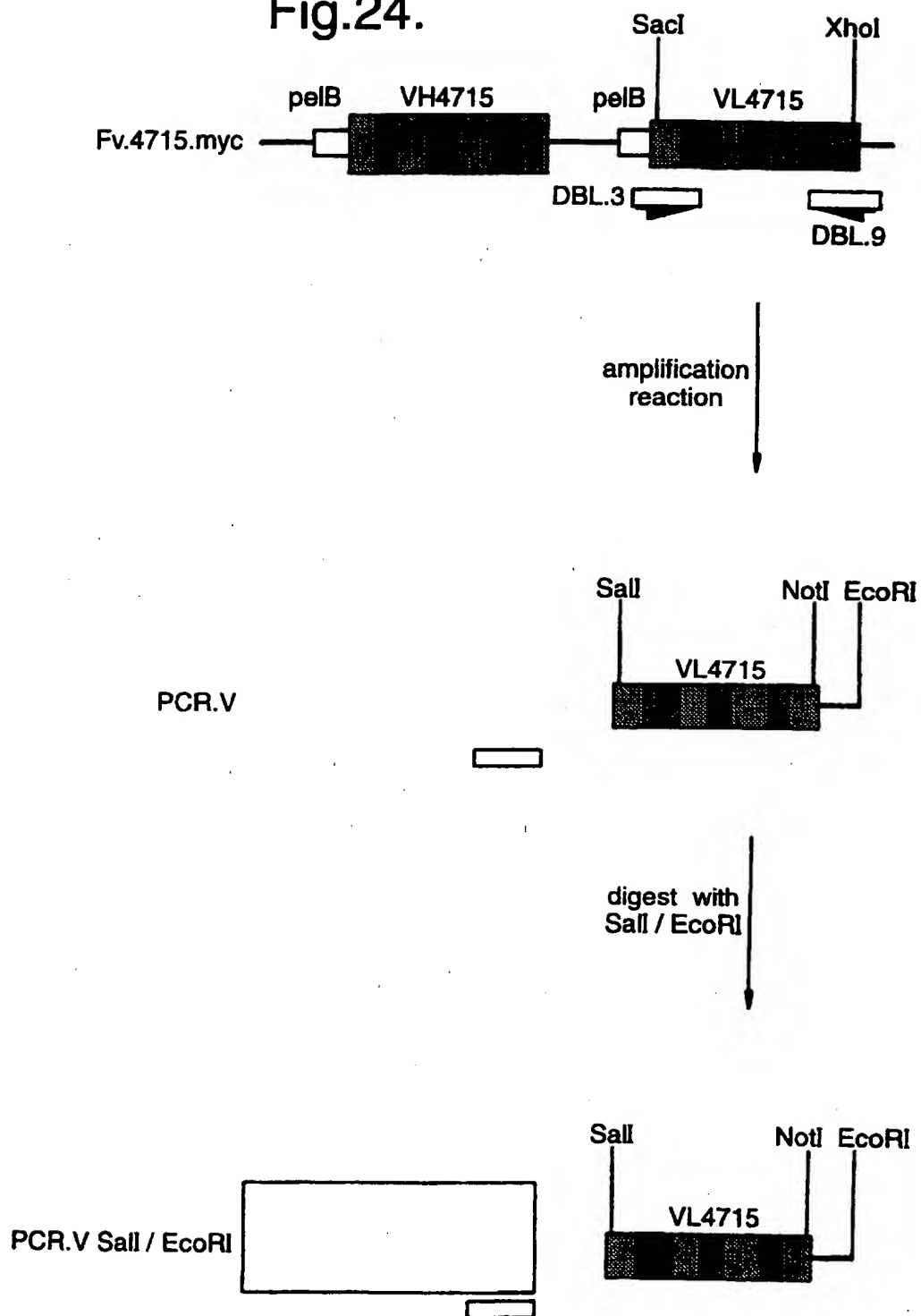
30/35

Fig.23.



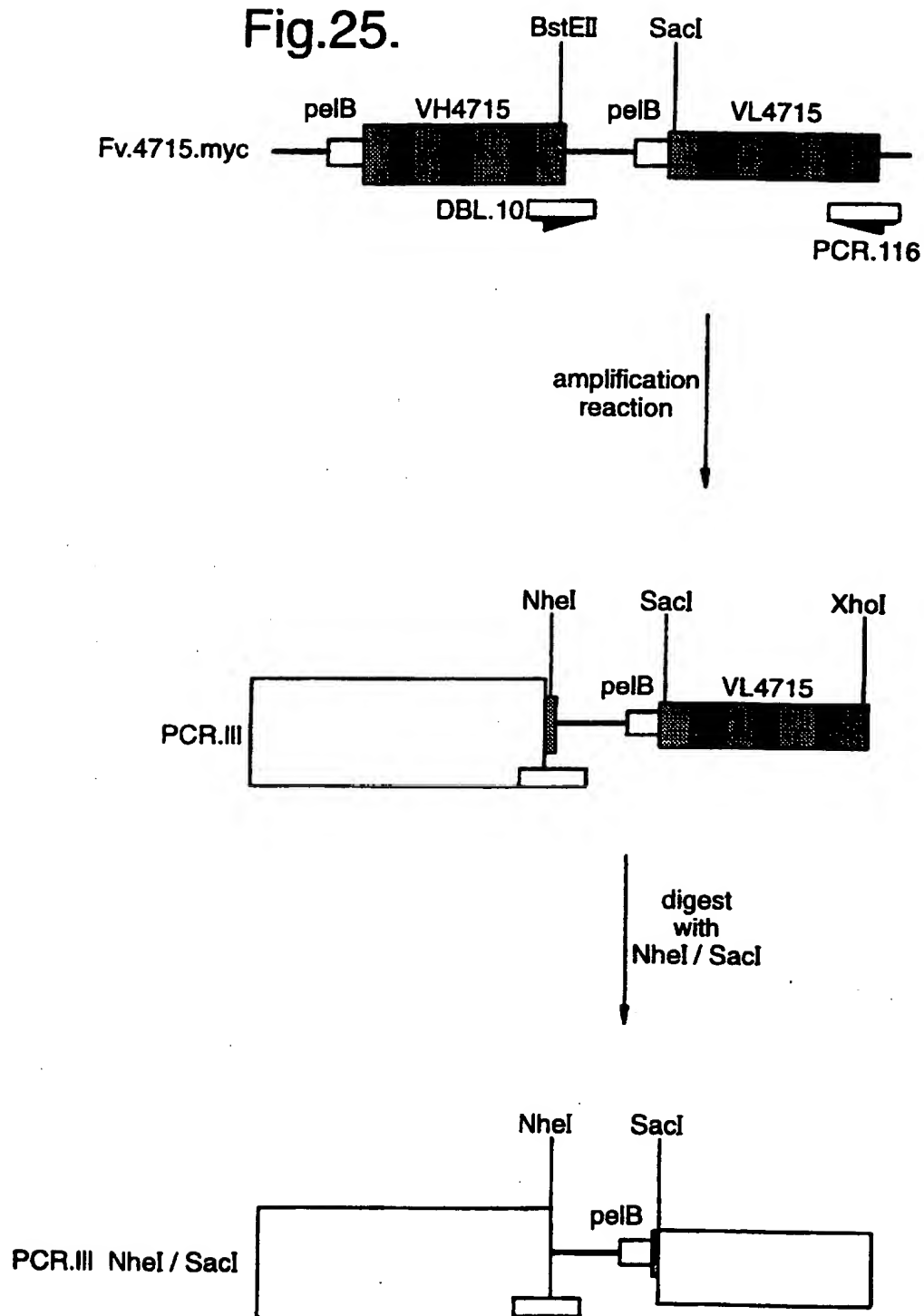
31/35

Fig.24.



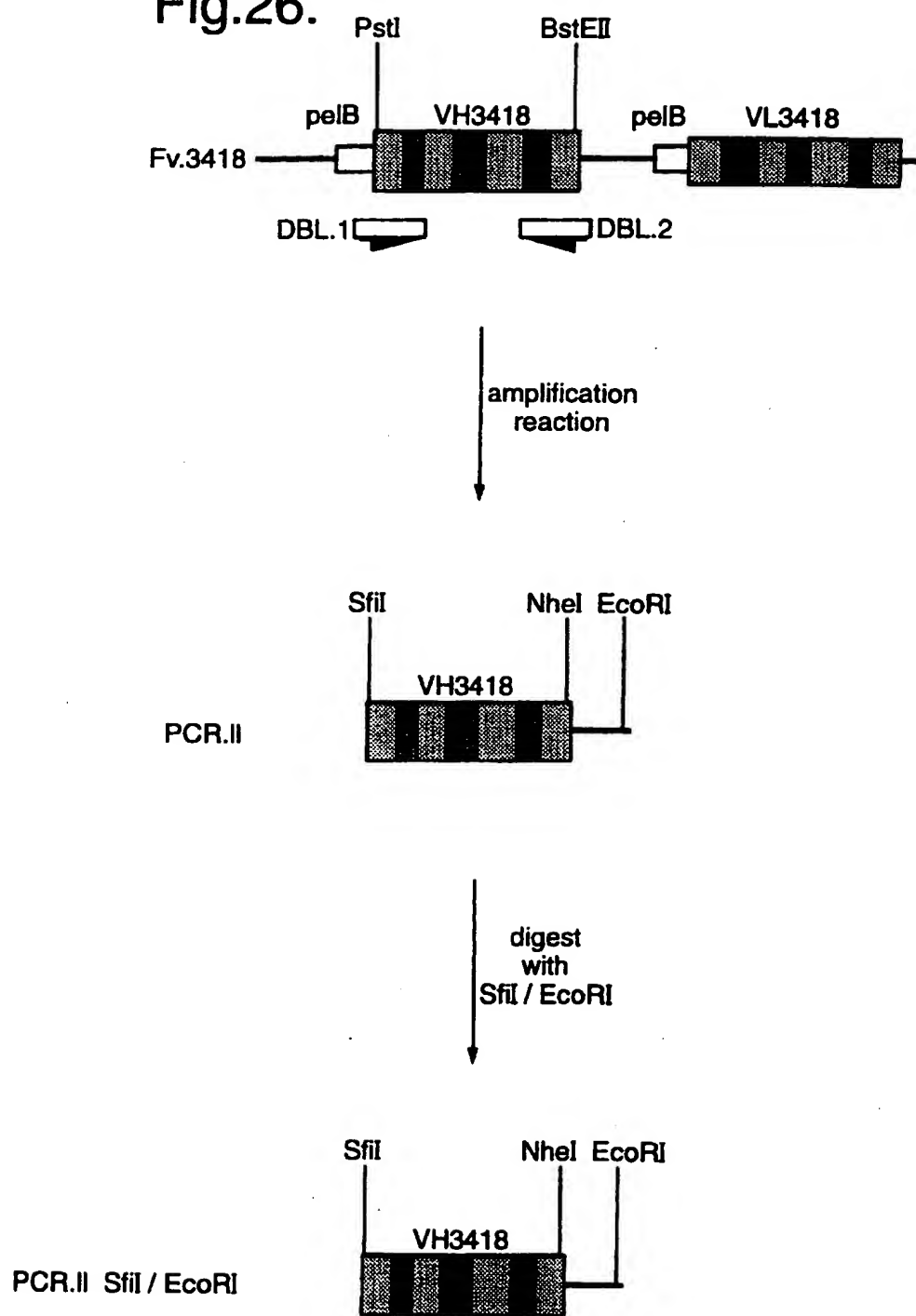
32/35

Fig.25.



33/35

Fig.26.



34/35

SEQ IDNO.

- 5 PCR.51 : 5' AGG T(C/G) (A/C) A(C/A)C TGC AG(C/G) AGT C(A/T)G G  
3'
- 6 PCR.89 : 5' TGA GGA GAC GGT GAC CGT GGT CCC TTG GCC CC 3'
- 7 PCR.90 : 5' GAC ATT GAG CTC ACC CAG TCT CCA 3'
- 8 PCR.116 : 5' GTT AGA TCT CGA GCT TGG TCC C 3'
- 9 DBL.1 : 5' CAG GAT CCG GCC GGT TCG GCC CAG GTC CAG CTG CAA  
CAG TCA GGA '3
- 10 DBL.2 : 5' CTA CAT GAA TTC GCT AGC TTA TTA TGA GGA GAC GGT  
GAC GGT GGT CCC TTG GC '3
- 11 DBL.3 : 5' ATT GGA GTC GAC ATC GAA CTC ACT CAG TCT CCA TTC  
TCC 3'
- 12 DBL.4 : 5' CGA ATT CGG ATC CCC GTT TGA TTT CGA GCT TGG TCC '3
- 13 DBL.5 : 5' GAG CGC GAG CTC GGC CGA ACC GGC CGA TCC GCC ACC  
GCC AGA GCC '3
- 14 DBL.6 : 5' AAT GTC GAA TTC GTC GAC TCC GCC ACC GCC AGA GCC '3
- 15 Euka.1 : 5' AGC TTC TAG ACC ACC ATG GAA AAC TGC AGA GCT CAA  
AAG CTA GCG CGG CGG CTC TAG '3
- 16 Euka.2 : 5' AAT TCT AGA GCG GCC GCG CTA GCT TTT GAG CTC TGC  
AGT TTT CCA TGG TGG TCT AGA '3
- 17 ML.1 : 5' ACG GGT GAG CTC GAT GTC GGA GTG GAC ACC TGT GGA  
GAG A '3
- 18 ML.2 : 5' GGA AAC AGC TAT GAC CAT GAT TAC '3
- 19 DBL.7 : 5' CAC CAT CTC CAG AGA CAA TGG CAA G 3'
- 20 DBL.8 : 5' ACC AAG CTC GAG ATC AAA CGG GG 3'
- 21 DBL.9 : 5' TGA AGT GAA TTC GCG GCC GCT TAT TAC CGT TTG ATT  
TCG AGC TTG GTC CC 3'
- 22 DBL.10 : 5' TAA TAA GCT AGC GGA GCT GCA TGG AAA TTC TAT TTC 3'

Table 1



35/35

Code	Expressed Antibody fragment
pSV.B	: VH4715-VH3418
pSV.D	: VH4715-VHlys
pSV.G	: VH3418-VHlys
pSV.K	: VH4715-VHlys-VH3418
pSV.M	: VHlys-VH4715-VH3418
pSV.N	: VL3418-VL4715.2t
pSV.P	: VLlys-VL4715.2t
pSV.S	: VLlys-VL3418.2t
pSV.V	: VLlys-VL4715-VL3418.2t
pSV.W	: VL3418-VLlys-VL4715.2t

Table 2

# INTERNATIONAL SEARCH REPORT

Internat. Application No  
PCT/EP 97/01609

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/13 C07K16/46 C07K16/00 C12N5/10 A61K39/395  
G01N33/577

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K A61K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 93 11161 A (ENZON, INC.) 10 June 1993 cited in the application see page 22, line 1 - line 10 see claims	1-15
A	WO 94 09131 A (SCOTGEN LTD.) 28 April 1994 cited in the application see claims see figures	1-15

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

14 August 1997

Date of mailing of the international search report

10.09.97

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+ 31-70) 340-3016

Authorized officer

Nooij, F

## INTERNATIONAL SEARCH REPORT

Internat Application No  
PCT/EP 97/01609

## C(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE USA, vol. 90, no. 14, 15 July 1993, WASHINGTON, DC, USA, pages 6444-6448, XP002014058 P. HOLLIGER ET AL.: "Diabodies": Small bivalent and bispecific antibody fragments." see the whole document ---	1-15
A	WO 94 13806 A (THE DOW CHEMICAL COMPANY) 23 June 1994 see figure 1 ---	1-15
A	WO 94 13804 A (CAMBRIDGE ANTIBODY TECHNOLOGY LTD. ET AL.) 23 June 1994 see page 31, line 10 - line 12 see figure 1 ---	1-15
T	WO 97 14719 A (UNILEVER) 24 April 1997 see the whole document -----	1-15

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/EP 97/01609

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 14  
because they relate to subject matter not required to be searched by this Authority, namely:  
Remark: Although claim(s) 14  
is(are) directed to a method of treatment of the human/animal  
body, the search has been carried out and based on the alleged  
effects of the compound/composition.
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such  
an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all  
searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment  
of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report  
covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is  
restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

Internat Application No

PCT/EP 97/01609

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9311161 A	10-06-93	AU 3178993 A CA 2122732 A EP 0617706 A JP 7501451 T	28-06-93 10-06-93 05-10-94 16-02-95
WO 9409131 A	28-04-94	AU 5283793 A CA 2146854 A GB 2286189 A JP 8505761 T	09-05-94 28-04-94 09-08-95 25-06-96
WO 9413806 A	23-06-94	AU 5747794 A CA 2117477 A EP 0628078 A JP 7503622 T	04-07-94 23-06-94 14-12-94 20-04-95
WO 9413804 A	23-06-94	AU 5654894 A CA 2150262 A EP 0672142 A JP 8504100 T AU 7621494 A CA 2169620 A EP 0720624 A WO 9508577 A	04-07-94 23-06-94 20-09-95 07-05-96 10-04-95 30-03-95 10-07-96 30-03-95
WO 9714719 A	24-04-97	AU 6873396 A	07-05-97